The PTE Academic and Outer Circle Students: Assessing Proficiency in English, Ownership of English, and Academic Performance at UK Universities

By

Roy George Wilson

Doctorate of Philosophy in
Applied Linguistics & English Language Teaching

At the

University of Warwick,
Department of Applied Linguistics

August 2016

Table of Contents

Acronyms and Abbreviations

CA – Cambridge Assessment (formerly UCLES)

CEFR – Common European Framework of Reference for Languages

CSELT – Country-Specific English Language Test

EC – Expanding Circle

EIL - English as an International Language

ELF – English as a Lingua Franca

ESRC – The Economic and Social Research Council

ETS – Educational Testing Service

GSE – Global Scale of English (Pearson's Score Range of 10-90)

IAET – International Academic English Test (IELTS, PTE Academic, TOEFL-iBT)

IELTS – International English Language Testing System

MESC – Majority English-Speaking Country

NNS – Non-Native Speaker

NS – Native Speaker

OC – Outer Circle

PICAE – Pearson International Corpus of Academic English

PTEA – Pearson Test of English Academic

SELT – Secure English Language Test

TOEFL – Test of English as a Foreign Language

UKBA – United Kingdom Border Agency

UKVI – United Kingdom Visas and Immigration

WAEC – West African Examination Council

List of Tables

List of Figures

Acknowledgements

I would first like to thank my wife Siu Kuen Wong (Wendy), for her patience and understanding in these last four years, and my two children, Kit and Leiwah for keeping me sane and happy while I researched and wrote this PhD. I cannot thank my parents enough, David and Zeena, for their unending support – both emotional and financial! And my brother Sam, sister Heloise, brother-in-law Altan, Arda and Talya.

I would like to thank my supervisors, Dr Claudia Harsch, and Dr Neil Murray, for their tireless feedback, patience and dedicated guidance on this PhD journey.

Many thanks also to my friends: To Erkan Kulekci and Zuleyha Unlu (now both Drs!) for their friendship and conversations during my infrequent trips to Warwick over the past four years; to Dr Marcus Simmgen for his advice following my upgrade panel; Dr Michael Howard for our many coffee chats; and Sam Pascoe for being a good friend.

I would also like to thank Dr Fiona English for starting me off in linguistics during her inspirational and innovative MA course at London Metropolitan University (2007-2008), and special thanks go to Dr Parvaneh Tavakoli who first interested me in language testing and assessment on the module at London Metropolitan University.

I dedicate this PhD to all my friends and relatives around the world from all circles of English and beyond!

Declaration

I declare that this research was conducted by myself and that all the work in this thesis is my own, and has not been used or published before. I also confirm that this thesis has not been previously submitted for a degree at this or any other university.

Roy George Wilson

Abstract

This mixed methods study explores the relationship of an academic English language proficiency test - the Pearson Test of English Academic (PTEA) - to the academic performance of its test takers at university. The particular focus is on the English language proficiency and academic performance of students from the "outer circle" (Kachru, 1985), many of whom have an accompanying ownership of English (Norton, 1997, Widdowson, 1994; Higgins, 2003). The implications of this proficiency and ownership of English for admissions, test performance and academic performance, are explored in the two strands of the study.

The first strand uses mixed methods including statistical analysis of a large data set of PTEA test scores; analysis of university admissions policy documents; and thematic analysis of interview and survey data. The second strand of the research uses interviews (tutorials) to investigate the interpretability of the test for four individuals from the outer circle (Anglophone West Africa), in particular, looking at what can be inferred from the PTEA score profiles about their English language proficiency at university. The strand investigates whether the test served any purpose for the evaluation of these four individuals' English language proficiency.

The study indicates that there are some statistically significant differences in the proficiency of test takers from the outer circle as compared to the expanding circle (as expressed through mean test scores) according to nationality sub-groups and L1s. Ownership of English is a theme manifested in the educational and social background of the survey and case study participants emerging from thematic analysis of the data.

Regarding inferences from test scores, the PTEA score profiles for the four cases do, to some extent, match their actual experiences at university regarding linguistic difficulties encountered on their courses. The case study also reveals other important factors in academic performance which are related to language proficiency but are also part of the adjustment to university experienced by all students including acquisition of academic skills and academic literacy.

The study makes a contribution to the study of World Englishes and language testing, in particular the legitimacy of the English of test takers from the outer circle and how they are viewed and processed in the admissions system to UK universities.

Introduction

This is a collaborative study jointly funded by the Economic and Social Research Council (ESRC) of the UK, the University of Warwick and Pearson. It is a mixed-methods study of an academic English language proficiency test - the Pearson Test of English Academic (PTEA). The study's central focus is the relationship of the outer circle test takers' proficiency in, and ownership of English to their PTEA test performance and subsequent academic performance and how students from this conceptual grouping should be assessed in terms of their English language proficiency for university entrance.

The first strand investigates the legitimacy of outer circle students as "native users" (Davies, 2013) of English in the context of admissions to university in the UK. Many outer circle students include those individuals for whom English is not a foreign or second language, and as such, they may be considered as being already proficient to a level of English sufficient for instruction in the medium of English.

As part of this investigation into the outer circle student and academic English tests, in the second strand of the research there is a focus on the inferences that can be made from test scores (predictive validity) of the test for the four outer circle case study participants. Predictive validity seeks to ascertain the existence and nature of a relationship between test scores used for university entrance, and the subsequent performance of students on their chosen courses in academia. For language testing researchers, the concern of predictive validity is whether a proficiency test produces scores that allow us to anticipate, infer or 'predict' whether a student will cope with the demands of the language of instruction in the post-test situation (in the domain of academia). For students whose first language is not English, an academic English language test is often required in order to demonstrate sufficient English language proficiency for study at an English-medium university. For the

test takers in this study there was also a question as to whether it was appropriate that they were asked to take the PTEA given the fact that they had already been schooled in the medium of English in their outer circle contexts.

It is necessary to state why it is important to look at whether inferences can be made about a test taker's score (predictive validity). Firstly, it is important to discover if students are adversely affected by being admitted onto a course with a certain test score. This is to prevent "wastage" in educational systems and to prevent students being "set up" for failure (Elder, 1993, p. 88). At one extreme this wastage refers to students dropping out of courses due to an inability to cope with the language demands of the course. At the very least it often means students needing to take language support classes in order to cope with a deficit in their language level. In this way, wastage also refers to "cost" (Banerjee, 2003) - any extra time and effort spent by staff and students in assisting the student to cope with their studies on the course. Secondly, information from studies which look at aspects of predictive validity can be used to provide the right kind of support for students at university. This is particularly helpful if the academic language test has a breakdown in sub-skills such as that provided by the PTEA, which might help to identify a particular skill in which students are deficient, allowing for language support to be focused in that area. Ultimately, we need to discover whether the tests are fit for their main purpose, which is to provide a snapshot of a student's proficiency in order to determine whether students are linguistically ready for the demands of tertiary education delivered through the medium of English.

This study focuses on the test as it relates to outer circle students, specifically exploring their proficiency in English for the domain of higher education in the UK. The outer circle (Kachru, 1985) refers to countries of the former British Empire and current Commonwealth:

countries with English as an official language, lingua franca and often a current or historical use of English as a medium of instruction in education and a wide range of other domains. Examples include Nigeria, India, Singapore, South Africa and Kenya. The outer circle also includes former and current territories of the USA[1], for example the Philippines. The inner circle (ibid) describes countries with the majority of their populations using English as a first language; specifically the UK, the USA, Australia, Canada, New Zealand and the Republic of Ireland.

Using Kachru's model or paradigm of World Englishes (1985), the major focus of this thesis is that a significant characteristic of English proficiency in the outer circle stems from the legacy of Empire in the existence and use of English in education systems and a variety of domains in domestic and public life. This historical legacy has shaped societies and produced large numbers of students from regions such as Anglophone West Africa who display a significant "ownership" of English (Gupta, 2006; Higgins, 2003; Norton, 1997; Widdowson, 1994, 1997) in both a localised and internationally intelligible form. In terms of proficiency, many students from these regions thus possess a "functional nativeness" (Kachru, 1998) and can be considered as de facto native speakers or at the very least highly proficient second language users of English, what Davies terms "native users" (2013). This conceptual grouping of outer circle test takers - their ownership of, and proficiency in English - has been overlooked in past studies of test performance and academic performance and so it is important to explore how academic English tests measure outer circle English proficiency and how these students perform in academia in the inner circle.

[1] Often as a legacy of American involvement in the Pacific Ocean in the Second World War.

In order to explore these issues, a qualitative and quantitative (mixed methods) approach has the potential to reveal interesting data including: what trends can be gauged in terms of test performance according to variables such as nationality and L1; what variables can affect the academic performance of students; and whether the PTEA test scores can allow inferences to be drawn about anticipating students' linguistic difficulties during their studies. This leads to the wider outcome of the study, which is to provide better support for international students by not seeing them as a homogenous group in terms of linguistic proficiency and to provide an insight into the linguistic challenges and other issues in academia that these students may face.

The two research aims in the two strands of the study are therefore concerned with: (a) exploring the particular proficiency of outer circle students for academia in the inner circle (focusing on PTEA mean scores, and UK university admissions policies), and (b); establishing whether any inferences can be made from the PTEA score report for the linguistic and academic performance of outer circle students in the domain of academia.

In the first strand of the research I use a large set of PTEA test scores to explore what the test 'says' about the English proficiency of outer circle test takers from various countries and regions and with various L1s, together with UK universities' admissions data on accepted previous secondary or high school English qualifications from outer circle countries of origin. In this strand I also collect qualitative data from interviews with Pearson country representatives and an outer circle student survey with a view to shedding light on proficiency in English and ownership of English. I then use largely qualitative case study data in the second strand in order to illuminate how individual outer circle students cope in their studies and to what extent language proficiency is a factor in their performance, relative to

their PTEA scores. I do this by comparing the PTEA score reports (appendix 1) of students against their actual experience and performance at university by looking at data elicited via semi-structured interviews (tutorials) with students from two Anglophone West African countries – Nigeria and Ghana. In doing so, it is also the aim of this study to provide an alternative model to more quantitative approaches to determining predictive validity.

Chapter 1 provides background information in order to set the scene for the study. I briefly present the Kachruvian model and paradigm of World Englishes, and give an overview of the admissions situation regarding language proficiency in the UK Higher Education system before profiling the PTEA. This background provides the context in which my study is located before the issues are discussed in more depth in the main literature review which is split into two sections within chapter 2 to deal with the two strands of the mixed methods study. Discussion of the literature around the first strand involves looking at the concept of the outer circle test taker and a critique of the Kachruvian paradigm of the circles of English, in relation to the literature on proficiency and the connection between ownership and proficiency. I also discuss issues around UK university admissions practices regarding outer circle students and their proficiency and a discussion on academic proficiency incorporating academic literacy. Discussion around the second strand first looks at the literature to do with validity and predictive validity (including the PTEA claims for predictive validity) and academic performance and the role of proficiency in academic performance. I then review past studies that have attempted to investigate readiness for, and performance in, academia and which have measured the predictive validity of tests and subsequent academic performance of students at university. I analyse and critically discuss the themes, methodology and results that have emerged from these studies and how they can inform my study.

Both chapters 1 and 2 outline the issues, context and my conceptual framework leading to articulation of the research gap and research questions at the end of chapter 2. The methodology chapter (chapter 3) then outlines the design of the study in relation to the research questions. This includes the design and rationale behind the mixed methods approach, and the data collection strategy and general analytical approaches employed to answer my research questions. Chapter 4 is a description of the data collection, analytical process and findings for the specific data sets followed by discussion and interpretation of the findings in chapter 5 where I involve the data from the two strands interpreting the findings from both to answer the research questions before concluding the study in chapter 6.

Chapter 1: Background

1.1 Kachru's Model of World Englishes

It is my intention here to introduce some key concepts in the study as a background to the literature review starting with Kachru's model of World Englishes.

All varieties of English can be termed "World Englishes" since they are part of a bigger "family" of Englishes. The concept of the outer circle has its origins within the concept of World Englishes emerging largely from the work of Braj Kachru (1985; 1992) who classified speakers of English throughout the world into three groupings of English based on nationality – the inner, outer and expanding circles (Figure 1) (Kachru, 1992, p. 356).

The inner circle comprises countries that have a majority of English speakers as "mother tongue" speakers traditionally considered "native speakers" of English. These include the USA, the UK, Australia, New Zealand, Canada and Ireland. Certainly in the testing sense, these countries are seen as norm-setting because they produce the norms in vocabulary, pronunciation and other aspects of language on which the tests are based (particularly the USA, the UK and Australia) (Ackermann, De Jong, Kilgarriff, & Tugwell, 2010; Lowenberg, 1993, 2000; Obaidul Hamid, 2014).

The outer circle comprises Englishes in Asian, Pacific, Caribbean and African countries which, along with inner circle countries, were also once part of the British Empire, from Ascension Island to Zimbabwe. Most of these countries and territories now comprise the Commonwealth where English is often a national or official language, a lingua franca or second language for many and a native language for a certain segment of the population,

usually affluent and urban, (Crystal, 2003, pp. 62-65; Graddol, 2006; Gupta, 2006). In many of these states and territories, English remains the language of government, commerce and indeed national education particularly in highly multilingual post-colonial states where (in terms of policy) English was often seen as a pragmatic and politically neutral choice for the purposes of nation building, national unity and education. Particular examples include India, Singapore, Nigeria, Kenya, Uganda, Zambia, Ghana, and Sierra Leone among many others. It is interesting to note that Kachru defines the Anglophone Caribbean nations such as Jamaica, as outer circle (Kachru, 1996, p. 137) - perhaps because of their English-based Creoles - but in terms of university admissions and academic English tests, countries in the Anglophone Caribbean are considered to be majority English-speaking and therefore their students are exempt from taking tests such as the PTEA.

The expanding circle meanwhile, contains countries such as Japan, China, Indonesia, Brazil and Turkey. Generally speaking, these countries did not have a deep historical relationship with the UK (in terms of being systematically colonised), with the result that English was not introduced and institutionalised in the same way as in those countries of the outer circle. The conceptualisation of learning and teaching English in these countries is more likely to be one of "English as a foreign language" (EFL). These countries are seen as norm-dependent because they depend on the inner circle norms as models for standards of pedagogy and assessment although this is now being challenged by scholars, notably by researchers in the developing area of "English as a Lingua Franca" (ELF) and "English as an International Language" (EIL) (Canagarajah, 2006; Jenkins, 2006, 2014; Seidlhofer, 2005).

Figure 1: The 'Circles' of World English

Adapted from Kachru (1992, p.356)

The Englishes in the outer circle can be given many names or labels. They are referred to in terms of chronological historical spread: Pakir, (1997) calls the outer circle the "second diaspora" of English as does Jenkins (2009) who differentiates these speakers from the "first diaspora" when English was spread by colonists and settlers to North America, and Australasia. However, because the circles represent the historical spread of English, there is also reference to the outer and expanding circle as the "third diaspora" (Kachru, Kachru, & Nelson, 2006)[2] because the earliest historical first diaspora would refer to the spread of English from England to Wales, Scotland and Ireland, - and second from the British Isles to the colonist and settler communities of North America and Australasia as above.

This historical spread of English has led to new varieties of Englishes (NVEs) although some of these "new Englishes" (Graddol, 2006, p. 84; Taylor, 2009, p. 143), have a long history in places such as West Africa and India (Kachru, 1986; Schmied, 1991; Schneider, 2007). However, they can be described as new in terms of being in the process of becoming established and recognised. For example, Crystal describes outer circle countries as places

[2] In this study B.B. Kachru will be referred to as 'Kachru'. Yamuna Kachru is refered to as 'Kachru' in this citation only.

where second language standards are "developing" and where local varieties of English such as that used in Singapore and Malaysia may gain more status as speakers of these varieties gain more influence in the world (Crystal, 2012). Kachru also uses this term when referring to these varieties as norm-developing because they are still being debated and established as recognised varieties of English. As they are norm-developing, one key characteristic of these varieties is that they have yet to be, or are in the process of being, codified or standardised (Anderson, 2009; Pakir, 1997) and thus cannot reliably be used in language tests as corpora.

From a political-historical perspective, these new Englishes have been labelled according to their political status for example, "post-colonial" (Schneider, 2007) denoting the development of these Englishes since the independence movements in the 1950s and 1960s and the creation of "new" post-colonial states in the Caribbean, Asia, Africa and the Pacific:

> "It was during this period that post-Imperial Englishes were being gradually institutionalized in the language policies of the changed political, educational, and ideological contexts of what were earlier the colonies of the UK and the USA" (Kachru, 1997, p. 66).

The Imperial British era has thus created a proliferation of "nativised" outer circle Englishes between the "native" varieties of the inner circle and the "non-native" varieties in the expanding circle (Higgins, 2003; Rubdy, McKay, Alsagoff, & Bokhorst-Heng, 2008), that in language testing terms are still deemed "non-native" and therefore non-standard. The history represented by Kachru's model has relevance for university admissions because admissions policies are based on nationality as an assumption of English proficiency (UK Visas and Immigration, 2016b, p. 36) and group all citizens of a country together in that respect. Therefore, UK government and university policies regarding admissions and language proficiency to some extent reflect Kachru's model. International students – from both the outer and expanding circles - are often obliged to take an academic English test in order to

demonstrate that they have sufficient proficiency in English for the domain of university. However, in this thesis it is my stance that the model represents an inheritance of English for nations and individuals which means that many students from the outer circle already have a combined ownership of, and proficiency in English - often as a first language or a fluent second language in an extensive variety of domains (Kachru, 1998; Lowenberg, 1993, 2000) including communicative competence through education (Davies, 2013). This then raises questions such as whether it is legitimate or valid to subject all outer circle students to language proficiency tests even when they have already been socialised and educated in English in their home outer circle countries.

Thus, given their background in, and relationship to English, it is important to investigate how outer circle students perform on academic English tests and subsequently in academia. The PTEA is one of these tests taken by many outer and expanding circle students in order to demonstrate their English proficiency for university study in the UK and other countries. This brief description of the circles model and its meaning as regards proficiency and ownership of English provides a context or theoretical framework for the study. This is discussed and critiqued in more detail in chapter two. I now outline the university admissions policies of the UK government and universities towards outer circle students and then profile the PTEA.

1.2 Determining Academic English Proficiency in the UK - Policies

1.2.1 The UK Government Policy and University Autonomy

Large numbers of students from across the world come to the UK each year to pursue higher education (HE). Many of these students are required to demonstrate proof of their English

language proficiency in order to satisfy the admissions criteria of their particular institution - and ultimately UK Visas and Immigration (part of the British Home Office[3]) - that they have sufficient levels of English proficiency to be able to cope with the linguistic demands of their respective academic courses. There are many language proficiency tests, both domestic and international, that can provide this "proof" or evidence as to a student's linguistic proficiency. These tests claim to assess proficiency in "academic English" in order to enable institutions to make a decision about whether to admit a student to a particular course at that institution. These decisions are based on inferences made about test scores and implicitly claim to be predictive in nature. In the UK, these tests include the well-established IELTS, TOEFL-iBT and now, the PTEA (launched in 2009). Other tests are used such as qualifications from Cambridge Assessment: Cambridge English First (FCE); Cambridge English Advanced (CAE); and Cambridge English Proficiency (CPE), and tests from Trinity College London, City & Guilds among others. For the purpose of this study, these tests are referred to collectively as "international academic English tests" (IAETs) some of which may be more academic in content than others.

From the point of view of the UKVI, international students from the inner circle are considered to be from a "majority English-speaking country" or MESC (UK Visas and Immigration, 2016b, pp. 34-35) (appendix 2). They are usually exempt from taking IAETs such as PTEA if they can provide alternative English qualifications from their domestic school system as proof of English language proficiency. Anglophone Caribbean nations (who are considered outer circle by Kachru) are also exempt from doing IAETs because according to the UKVI, these countries are categorised alongside inner circle countries as MESCs (ibid). The situation is different for applicants to UK universities from the rest of the countries classified in the outer and expanding circles. All outer and expanding circle

[3] The specific department once called the UK Border Agency (UKBA), is now called 'UK Visas and Immigration'

international students are considered potential takers of IAETs whether they are from outer circle Commonwealth countries such as India, Malaysia, Singapore, India, Nigeria, Ghana or Kenya or expanding circle non-Commonwealth countries, such as Japan, South Korea, Russia, Brazil, Turkey or China. This is regardless of whether the student has been educated in the medium of English or whether English is a first, second or foreign language, or a lingua franca.

Officially in the UK, the English proficiency of these students should be assessed by an IAET additionally approved by the UKVI as "secure". An IAET given "Secure English Language Test" status (SELT) by the UKVI is recognised by this body as being official proof of linguistic ability (UK Visas and Immigration, 2016a). The comprehensive list of these SELTs is available in a document from the British Home Office (ibid). IAETs given SELT status are periodically reviewed by the government with approval of an IAET as a SELT being subject to withdrawal at any time. At the time of writing (summer 2016), PTEA (Pearson) and TOEFL-iBT - produced by Educational Testing Service (ETS) in the USA - which have been approved as SELTs in the past are no longer on the approved SELT list but may be added again in the future. Trinity exams and IELTS remain as SELTs (ibid). This does not mean that the PTEA test is no longer an *academic* English test - it remains so, and universities are free to select it as a means of assessing the linguistic readiness of students for university level study (Pearson, 2015).

The UKVI bestows "Tier 4" status on universities giving them the right to sponsor students for visas and thus the ability to recruit international students because they have a "trusted sponsor" status (UK Visas and Immigration, 2016b). This means that universities, as Higher Education Institutions (HEIs), are currently free to use whatever evidence they deem suitable

as proof of this proficiency, either by using SELTs, non-SELTs or other means such as interviews or internal university tests.

However, sole use of outer circle students' home qualifications as evidence of proficiency may be seen as a risk by some universities and contributes towards a politically and financially influenced admissions system in which there are tensions between, on the one hand encouraging international student recruitment through giving students "the benefit of the doubt", and on the other a strict adherence to linguistic thresholds set by the UKVI in order to maintain official linguistic entry standards (Banerjee, 2003, pp. 369-370; Bayliss & Ingram, 2006; Edwards, Ran, & Li, 2007, p. 388; Feast, 2002, p. 84; Ingram & Bayliss, 2007, p. 54) . It is into this political market of academic English tests and SELTs that the PTEA has entered, a market that includes test takers from the outer and expanding circle who possess a wide variety of Englishes and experiences with English.

1.2.2 Use of Alternative English Language Qualifications

Despite the seemingly clear-cut "rules" on who has to actually take a test, and what test is acceptable for proof of English language proficiency, many universities in practice actually use older and more established ways to assess an applicant's English language proficiency than use of SELTs. The evidence of proficiency collected by a university must be stated in a letter called a "confirmation of acceptance for studies" letter (CAS) (UK Visas and Immigration, 2016b, p. 14) which universities issue to students to secure their student visas. This evidence can take the form of, for example, a university's own internal "in-house" test; a recruitment interview; or a student's qualifications in the medium of English from their domestic educational system - a "country-specific English language test" (CSELT – see

appendix 3). The individual universities decide upon this practice "If your sponsor is an HEI, we will allow them to choose how to assess your English language ability" (ibid, p.33), but the evidence must indicate that the student has reached a level that the UKVI approves of (see next section below).

Some university websites display a list of the CSELT qualifications that are acceptable as alternative proof of English proficiency (Swansea University, 2014; University of Leicester, 2016; University of Liverpool, 2014). These CSELT qualifications range from the Botswana Certificate of Secondary Education to the Zambian School Certificate, and most were inherited and developed from the British system at the end of Empire (Read, 2015, p. 3). The position of students in outer circle countries is that their English-medium high or secondary school CSELT *can* be accepted as proof of English language proficiency by universities. Whilst this practice does at least indicate an acknowledgement of the prior educational experiences of outer circle students in the medium of English, universities must tread a fine line between on the one hand, using their own defined indices of linguistic proficiency and on the other, accepting students based on criteria that they, *and* the Home Office deem acceptable. In light of the latter, it is not uncommon for many students from the outer circle, to be required to take a SELT irrespective of the level of proficiency or background they already have in English.

Regardless of the assessment criteria – whether the student has done a SELT, undergone a university interview, taken an internal university proficiency test or been accepted on the basis of their CSELT - the government sets the actual linguistic threshold level for entry into university. The UKVI stipulates that the linguistic level at which a student is accepted for

university (using whatever assessment method) should be at a level of English proficiency equivalent to level "B2" of the Common European Framework of Reference (CEFR).

1.2.3 The UKVI "B2" Level for University Entrance

The threshold level of proficiency for international academic English tests is set by the UKVI at B2 of the CEFR (UK Visas and Immigration, 2016b, pp. 33-34). The CEFR (Council of Europe, 2001) has a six-tiered set of competencies for assessing proficiency from levels A1 (lowest level) to C2 (highest level). Each band in the CEFR has associated descriptors that describe what that level of competency or proficiency might be imagined as being, in the performance of a student, and can be seen in the global scale of the CEFR (Council of Europe, 2001, p.5). It should, however, be borne in mind that the UKVI set the threshold with a focus on immigration which is not necessarily in line with university concerns:

> The UK Border Agency has set English language entry requirements that they feel are necessary to achieve the goals of immigration policy. Your university needs to set English language entry requirements that match its own goals for international recruitment and international student success. (Howell et al., 2012, p. 6)

Although universities can set their own levels for entry (ranging from A2 to above B2), the UKVI will not issue visas unless a B2 level for each student is *demonstrated* and *evidenced* by the universities (Howell et al., 2012, p. 7).

Level B2 is conceptualised by Pearson for the PTEA in the form of descriptors taken from the PTEA Score Guide document which aligns test scores to the CEFR (Pearson, 2012a, p. 47). Level B2 is aligned to a score range of 59-75 on Pearson's scale for measuring proficiency on

their tests – the "Global Scale of English" (GSE) (appendix 4). In appendix 4 the text on the right side in the table describes the rationale and meaning behind the B2 level for academia. The descriptors cover all four language skills – the three sentences beginning with "can" refer to reading, speaking & listening, and writing respectively. These imply that anyone scoring at this level would be expected to cope with their university studies *linguistically* and not encounter problems due to low language proficiency. In this way, test scores aligned to B2 constitute a type of "passing score" for university entrance which represent "an appropriate performance standard" (Kane, 1994, p. 426).

In order to justify this level for university, the testing organisation Pearson does however, make the following statement about the level *below* B2 in their academic score guide document (appendix 4); namely, level "B1":

> B1 is insufficient for full academic level participation in language activities. A student at this level could 'get by' in everyday situations independently. To be successful in communication in university settings, additional English language courses are required. (Pearson, 2012a, p.41, 2012b, p.4)

In the CEFR levels there is also a B1+ level indicating that "scores in this range predict success on the easiest tasks at B2" (2012a, p.40). In this sense, a student at B1+ is almost at B2 whilst B1 is deemed an insufficient level for academia and indicates a need for additional support. These levels of B1 and B1+ can be the basis for accepting students onto university pre-sessional courses, which are language courses that universities accept students on before they start their degree courses which in theory allow students to build up to B2 level before starting their degree proper.

I now outline the features of the PTEA, which, as a test designed to measure academic English proficiency for university study, is the subject of this study.

1.3 The Pearson Test of English Academic (PTEA)

The PTEA has been in use since 2009 and competes with other well-established academic language tests such as IELTS and TOEFL-iBT in providing tests for assessment of English proficiency for university admissions. Test specifications outline what is tested, how and who is tested – in short, "what theories of language and proficiency underpin the test" (Alderson, Clapham, & Wall, 1995, p. 10). These can be seen in appendix 5 and a summary of these is given below:

1. The PTEA has been developed from a corpus of academic English comprising inner circle countries of the UK, the USA, Australia, Canada and New Zealand (Ackermann et al., 2010) and thus acknowledges a diversity in standard English but one which is restricted to the inner circle.

2. The PTEA scores have a concordance with the language descriptors in the CEFR (2010a, 2012a, 2012b), (appendix 4). This alignment of the PTEA scores with the CEFR bands is to "ensure comparability and interpretability of test scores" (ibid, 2012a, p.38). The comparability is in reference to other tests and assessments while the interpretability refers to how the PTEA scores can be interpreted by stakeholders such as university admissions departments regarding the suitability of a student's English language proficiency for a course of study.

3. The PTEA takes an integrative skills approach to testing, which means that any combination of the four skills of speaking, listening, reading and writing may be tested in any

one test item or task. The test has 20 different item types providing a candidate with 70-91 items in the 3-hour test, which is delivered entirely through a computer (Pearson, 2012a, p.2).

4. The test (including the spoken and written components) is marked by computer and from this, a score report is produced giving candidates an overall score as well as ten sub-skill scores (see appendix 1). Because the test uses a computer marking system, there is a claim that the test has less rater bias and is more reliable than other tests (Pearson, 2012c, pp. 1-2).

5. The score report produced by Pearson for each test taker consists of an overall score between 10 and 90 on the Pearson GSE and a score on the same range for the four "communicative skills" of reading, writing, listening and speaking, as well as six further scores for the six "enabling skills" of grammar, oral fluency, spelling, pronunciation, vocabulary and written discourse (see appendix 6). The communicative skills are designed to inform university admissions decisions and the enabling skills are intended for the student's own language development and reference because they are not suitable for use in "high-stakes decision making" (Pearson, 2012a, p.43).

6. Regarding score thresholds for the purpose of university decision-making, Pearson advise the following:

- For undergraduate studies a minimum score between 51 and 61
- For postgraduate studies a minimum score between 57 and 61
- For MBA studies a minimum score between 59 and 69

(Pearson, 2012b, p. 3)

It should be noted that the Pearson recommended score for undergraduate study begins at 51, which is aligned to B1+ in the CEFR - below the UKVI threshold level of 59 (B2). Similarly

for postgraduate study the minimum score recommended by Pearson (57) is still below the B2 level set by the UKVI.

With regards the focus of this study on the English proficiency of the outer circle student, their ownership of the language, and the validity of the test for these students, it is necessary to look at two specific PTEA claims that align with the two strands of this study:

1. Firstly the claim in the test specifications that the "intended participants" of the PTEA are "non-native speakers" (Pearson, 2010b, p. 1) and "learners of English as a second or other language who are applying for admission to courses where English is the language of instruction…" (appendix 5).

2. Secondly the 'predictive power' of the scores articulated by the alignment with the CEFR:

> it can be concluded that the level B2 is required to be likely to function successfully in language exchange as one may encounter in higher education. Basically this implies that students who have attained level B2 in a foreign language which is used as the language of instruction and communication in an institution for tertiary education would not be disadvantaged significantly because of the language in comparison to students for whom that language is their first language (Pearson, 2010a, p. 2).

I explore this claim in the second strand of the study, gathering evidence of "language exchange" – the communication and interaction that occurs in tasks and situations in academia - and looking at whether the outer circle students scoring at B2 (as cases) were not disadvantaged significantly (although in this study I did not have access to inner circle students whose first language was English to compare with my participants). I discuss these two claims in more depth in the two strands of the literature review in the context of other literature.

In summary, only a small number of the English language tests in the world, even internationally recognised ones, are actually approved by the UK Home Office as "secure" (UK Visas and Immigration, 2016a). For any of the academic English tests however, we have also seen that the proficiency threshold entry level for Higher Education is set at B2 by the UKVI and that it is still the prerogative or "right" of British universities as to how exactly they ascertain the English language proficiency of a student who applies to study with them. In order to place the outer circle student into this admissions context I now move on to the literature review in order to explore the concept of the outer circle test taker and the issues surrounding their proficiency and ownership of English as the focus of the study.

Chapter 2: Literature Review

2.1 Research Strand 1 – The English Language Proficiency and Ownership of English of the Outer Circle Student.

2.1.1 Proficiency

I start this chapter with a claim from the PTEA literature to illustrate what the first strand of the research is investigating:

Claim 1: The target test taker

> "If you are a non-native speaker of English and need to demonstrate your academic English language ability, PTE Academic is the test for you" (Pearson, 2010b, p.1)

The focus of this strand, and indeed study, is largely to explore the assumption in this claim that the test taker will be a non-native speaker of English who needs to be assessed for their ability to function in an English-medium academic environment. In order to address this claim, I first look at the issue of ability or proficiency in language, introduce the concept of academic English proficiency as a type of English for Specific Purposes (ESP) and its relation to academic literacy. I then discuss proficiency in relation to the Kachruvian paradigm of World Englishes and the circles of Englishes model. I make the link between ownership and proficiency in English in more detail around the English "nativeness" of the outer circle student and the implications of this for language tests and university admissions. How these concepts relate to validity and academic performance in this study is covered in the second strand of the literature review.

A definition of proficiency in a language is hard to pin down – the common word indicates "a high degree of ability or skill in something" (Macmillan, 2007, p. 1184). However, in linguistics, the definition is still contested "the concept of language proficiency, its structure, and measurement continue to be controversial in language testing" (Nunan, 2009, p. 312). In terms of the PTEA, the test incorporates proficiency in terms of being able to communicate in an academic setting and thus a performance of language knowledge in a particular context. In other words, an individual may have knowledge of a language but stakeholders need to know whether an individual can communicate in a language in a particular place or context. This concept of proficiency can be traced to the notion of communicative competence (Hymes, 1972). Hymes proposed "(tacit) knowledge and (ability for) use" (1972, p. 282) - the notion that communicative competence is the knowledge of language and skills in order to interact socially with an emphasis on knowledge of the grammar of language but also knowledge of the appropriateness of language for specific contexts (ibid, p.277).

This was a significant development from Chomsky's distinction between competence as the knowledge of language (generative grammar) and performance as "the actual use of language in concrete situations" (Chomsky, 1965, p. 4). Chomsky described knowledge of a language as "generative grammar" (Chomsky, 1965, pp. 15-16) which is "a system of rules that in some explicit and well defined way assigns structural descriptions to sentences" (1965, p.8). He differentiated the knowledge and use of the language and stated that performance "cannot directly reflect competence" (ibid). Canale and Swain (1980) conceptualised communicative competence as composed of grammatical, sociolinguistic and strategic competence:

> A communicative approach must be based on and respond to the learner's communication needs. These needs must be specified with respect to grammatical competence (e.g. the levels of grammatical accuracy that are required in oral and written communication), sociolinguistic competence (e.g. needs relating to setting, topic, communicative functions), and strategic competence (e.g. the compensatory communication strategies to be used when there is a breakdown in one of the other competencies). (1980, p.27)

Grammatical competence is a "knowledge of the rules of grammar" and sociolinguistic competence is a "knowledge of the rules of language use" (ibid, p.6). They therefore make a distinction as with Hymes and Chomsky about knowing, and knowing how to do something although in their construct of communicative competence the idea of actual performance of this knowledge is also separate:

> "Communicative competence is to be distinguished from communicative performance, which is the realization of these competencies and their interaction in the actual production and comprehension of utterances (under general psychological constraints that are unique to performance)" (ibid, p.6).

Bachman & Palmer (1996) also developed a similar construct which they termed "language ability" (p.66-70) consisting of:

- Organisational language knowledge (grammatical and textual)
- Pragmatic language knowledge (functional and sociolinguistic)
- A component of strategic competence (metacognitive components and strategies)

For the purposes of this study I interpret communicative competence or ability as a way to articulate the proficiency tested for in IAETS such as PTEA because proficiency is a knowledge of the grammar, its acceptability and an awareness of appropriate use of language (sociocultural) as well as having the skill to perform this knowledge – "how well one can perform this knowledge in actual communication" (Canale, 1983, p.5). Canale and Swain state that this communicative competence should also have implications for language testing particularly as it relates to "what extent the learner is able to actually demonstrate this

knowledge in a meaningful communicative situation (performance)" (ibid p.34). This idea of performing in communicative contexts is explored in the second strand of this study.

Proficiency however, may also be of different types, Cummins (1979, 2008) differentiated between an oral or conversational proficiency – "basic interpersonal communication skills" (BICS) and what he termed a "cognitive academic language proficiency" (CALP). CALP is his notion that after the early stages of schooling, children develop an oral and written proficiency that is more reflective of the language needed to access schooling (Cummins, 2008, p.71). Cummins's notion of academic English was developed in the context of primary and secondary schooling in the United States but the theory is transferable to higher education in the sense that there is a need to think of proficienc(ies) for specific purposes.

Hulstijns (2011, p. 232) notes a similarity between Cummin's BICS and CALP in his conceptualisation of language proficiency (LP) as "basic language cognition" (BLC) and "higher language cognition" (HLC). BLC comprises the language that all native speakers share – a basic speaking and listening ability. HLC is where native speakers differ in proficiency from each other referring to use of written and spoken language, and discourse away from the everyday – schools and colleges, on the work floor and in leisure-time activities (ibid, p.231). This differentiation by Cummins and Hulstijn of a general English proficiency and one for academic purposes leads into the area of proficiency for a particular domain which in this study concerns the idea of English for Specific Purposes (ESP).

2.1.2 Proficiency for Academia

2.1.2.1 A General Academic Proficiency

In this study, the main conceptualisation of proficiency as it relates to the PTEA is a communicative competence for a specific context of use. A test taker of academic English tests is therefore assessed as to their ability to "do something specific in the language, for example proficiency in English to study in higher education in the UK…" (Davies et al, 1999, p.153). The ability to be proficient in an English for academia is realized in the area of English for Academic Purposes (EAP) which is a type of English for Specific Purposes (ESP), a concept concerned with the design of syllabus and materials oriented towards the communicative needs of the learner (Munby, 1978, p. 2). EAP is a form of English for communication and interaction in a specific context or domain. In applied linguistics "domain" refers to "the area of language use or the social correlate of a particular speech style or variety, normally marked by such features as topic of discourse, locale of discourse and relative status of participants" (Davies et al., 1999, pp. 49-50). In this study, the domain is academia and academic English in the inner circle, the situations and tasks that students will deal with in academia and the tutors and students with whom they will interact and communicate. Academic language tests such as IAETs have been developed to encompass communicative competence as a core concept in their design (Bachman, 1990, p.297) and focus on a very specific future domain of language use:

> "there is no place for universal proficiency in English-language testing anymore. Proficiency can be addressed meaningfully in only specific contexts and communities of communication in relation to the repertoire of codes, discourses and genres that are conventional for that context" (Canagarajah, 2006, p.241).

Thus, an academic English test is not a test of a student's proficiency in his or her particular country's variety of English, but rather a test of proficiency for academia.

In the light of the above brief discussion of proficiency the Pearson literature makes reference to the terms "proficiency" and "communication" in the context of academia - for example, the PTEA measures:

- "English language proficiency to ensure success and active participation in university and college level education" (Pearson, 2010b, p.1)
- "the English communication skills of international students in an academic environment" (Pearson, 2010a, p. 1)
- "academic English language ability" (Pearson, 2010b p.1)
- "academic English language proficiency" (De Jong & Zheng, 2011, p. 3)
- "the English that students will need to understand and produce in order to be successful in academic settings where English is the language of instruction" (Ackermann et al., 2010, p. 1).

The above statements place the test in the communicative competence tradition of proficiency - being able to measure actual language and communication skills needed for the context of university study and evidenced by use of such terms as "communication skills" and "active participation" in "academic settings" or "an academic environment".

It can also be asked as to whether there is one type of academic domain and therefore whether there should be one type of academic English proficiency. On an abstract level, a general concept of academic proficiency is conceptualised in the alignment with the specific competencies expressed at the level of B2 in the CEFR. There is a sense in which the construct of academic English or academic language is "taken for granted" and that it is the same construct no matter what the discipline is (Davies, 2007, p.74). Davies defines academic language proficiency as "the language of argument, of analysis and of explanation and reporting, in all cases not being specific to any particular academic area"' (2007 p.85).

Having a general academic English language proficiency or competence provides a foundation on which students need to build other competencies and skills in order to cope at university. It is clearly important to have a certain level of language proficiency in order to engage with university life. Fox (2004, p.438) mentions that the "premise" of a threshold of language proficiency is "essential for engagement with academic work, but insufficient on its own to ensure the quality of that engagement" (2004, p.438). In terms of proficiency, once students enter the university and their disciplines they will have to acquire another level of proficiency in terms of academic literacy.

2.1.2.2 Academic Literacy as Proficiency

Academic literacy is a concept that needs to be addressed in relation to discussion of proficiency because it extends the notion of a general academic English proficiency to one situated in the contexts, disciplines and practices of academia. This is important when interpreting the linguistic performance of students at university.

The PTEA is an indicator of general academic proficiency, but not of academic literacy for a particular discipline. However, some discussion of academic literacy is important here to connect with the section of conceptualising academic performance (section 2.2.2 below) where language proficiency is assessed in relation to communication and interaction in a number of situations and tasks in academia that can be understood with reference to "academic socialisation" and "academic literacies" (Lea & Street, 1998).

Academic proficiency is often concerned with academic writing. Lea and Street argue that "educational research into student writing in higher education has fallen into three main perspectives or models" (1998, p.158). These models comprise: study skills; academic

socialisation; and academic literacies. This is a conceptualisation of academic literacy arising from the writings on "new literacies studies" of the 1990s (Barton, 1994; Baynham, 1995; Street, 1984). Lea and Street argue that the study skills approach is based on a "deficit" (1998, p.157) model of writing - that students are deemed to be experiencing problems that can then be fixed. They state that study skills emphasises "surface features, grammar and spelling" (p.159) skills which are reflected in the PTEA score profile reflecting language proficiency. Lea and Street argue that the area of study skills has been seen as "atomised skills which students have to learn and which are then transferable to other contexts" (ibid, p.158). These can be equated to general or "generic skills" (Murray, 2010, p.58), which are often the focus of EAP courses at university, or university "guidelines" for students (Lea & Street, p.164).

This "study skill" approach is incorporated within the next term of "academic socialisation" - concerned with academic staff, including tutors, "inducting" students into the "new culture" of the academy (ibid, p.159). Lea & Street acknowledge criticism of this label as one which "appears to assume that the academy is a relatively homogenous culture, whose norms and practices have simply to be learnt to provide access to the whole institution" (1998, p.159). Murray (2010) notes that there is also a problem in terms of thinking of a homogenous concept of academic literacy "the study skills approach, then, takes a one-size-fits all view of academic literacy, dislocating its subject matter from particular disciplinary contexts" (201, p.59) and mentions that this is a common approach in EAP and other programmes. Wingate & Tribble (2012) state that EAP has tended to be placed in this area of academic literacy - "all categories of EAP are subsumed under the label 'academic socialisation'" (2012, p.487). In terms of academic acculturation a university is a "discourse community" (Spack, 1988) into which students are initiated but it is the disciplines themselves that induct the students

into their specific traditions and conventions therefore, Lea and Street also articulate the concept of "academic literacies" (1998, p.159). Academic literacies are concerned with "student writing and learning as issues at the level of epistemology and identities rather than skill of socialisation" (Lea & Street, 1998, p.159). They talk of "literacy demands of the curriculum as involving a variety of communicative practices, including genres, fields and disciplines" (ibid). This also includes a concept of "contested meaning" over issues such as what is meant by structuring a piece of writing for a particular discipline or genre of writing (Lea & Street, 1998, p.165-167, Murray, 2010, p.59). Students and tutors have different expectations of what constitutes writing for a discipline or genre and bring different cultural practices in their writing such as content and discourse in essays (English, 1999) and so academic literacy is about "the processes of meaning-making" and "student interpretations of what is involved in student writing" (Lea & Street, 1998, p.159). The connection with a general academic proficiency is that a certain level of general English language proficiency is needed in order to perform in these other literacy practices at university (Murray, 2010, p.58).

There are other views on academic literacy for example, traditionally, academic literacy was considered to simply include the abilities to read and to write in "page-bound" contexts (Cazden, Cope, Fairclough, & Gee, 1996; Spack, 1997). Although it is important to be able to read and write - skills still largely central to academic literacy and a measurement of performance - there is an argument that this traditional view of academic literacy is "restricted to formalized, monolingual, mono-cultural and rule-governed forms of language" (Cazden et al., 1996, p. 61) and now the term "literacy" is thought of in a more plural and holistic sense. In a plural sense, academic literacy is about "literacies" in that each discipline has "literacy practices" which "help define and differentiate that discipline" (Murray, 2010, p.58). An "academic literacies" approach to proficiency in academia is one in which a student

needs to "negotiate a multiplicity of discourses" (Cazden et al. 1996, p.61). Gibbons (2009) posits that academic English is context-specific and therefore specific to the discipline that the student is in. She talks about academic literacies as involving "learning to control new language" (2009, p.5), and in academia *in a particular subject*. This is certainly demanded of students as they grapple with the specific terminology and concepts that are essential to becoming conversant in their specialised subject, whether it be engineering, maths, literature, the sciences or indeed linguistics. Gibbons also therefore argues for *academic literacies* rather than academic literacy and the onus is on the student to "understand the language demands of their own subject" (2009, p.8). In a holistic sense, these discourses are social, academic and also cultural:

> Graduate students not only need to build interactive relationships with their teachers, thesis supervisors, and peers, and develop effective research strategies and good writing skills, they also need to adapt smoothly to the linguistic and social milieu of their host environment and to the culture of their academic departments and institutions (Braine, 2002, p. 60).

This suggests a relation to Bourdieu's notions of "cultural capital" and "habitus" (Bourdieu, 1973, 1986) in terms of what expectations and culture students bring to the academy and what cultures of writing and learning they are socialised into during their time at university. Similarly, Bayliss and Ingram (2006) listed some behaviours that were important for university study, behaviours that align with the concept of academic literacy as it relates to "discourses of power" (Lea & Street, 1998, p.158-159):

> Behaviours such as questioning or contradicting academic staff, participating in argumentative debates, applying critical thinking strategies, independently managing one's study regime and attending classes with other genders, age groups, social and cultural groups are all part of the Australian university experience (Bayliss & Ingram, 2006, p. 10)

All these articulations of academic literacy or literacies as part of the proficiency demands of academia means that predicting academic performance from inferences made from academic

English test scores is complex. The existence of academic literacies suggests that in order to perform on a course of study, a student needs to develop a type of proficiency beyond the "general" academic English proficiency tested for in IAETs. This is because the concept of performing academically goes beyond language proficiency skills or "surface features" and even language proficiency is very much discipline-dependent after a certain level - "academic success may require different levels of language skills at different campuses and for different academic majors" (Light et al, 1987, p.253).

Furthermore, regarding academic literacy, it is important to stress that acquiring academic literacies is important for all students at university, whether from the inner, outer or expanding circles, and regardless of whether they have English as a first, second or foreign language (Wingate & Tribble, 2012) or have taken an IAET. Neither students from English-speaking nor non English-speaking backgrounds will necessarily have been equipped with academic literacies prior to their arrival at university (Murray, 2010, p. 61). All students have to adapt to university and acquire a proficiency for academia beyond a general academic English threshold. From this perspective, a student may be highly proficient in a language and even be of an English speaking background but they will still need to acquire academic literacies in order to perform in academia.

Issues of proficiency as conceptualised by literacy are also discussed in relation to "aspects of academic performance" (section 2.2.2) because although the investigation of predictive validity and performance in this study is looking mainly at a linguistic performance at a level of general academic English, it also needs to take into account that students will encounter performance issues which are discipline-specific and based on skills and abilities other than

language proficiency such as that covered by notions of "academic socialisation" and "academic literacy".

Regarding proficiency as it relates to the outer circle, I now turn my attention to the Kachruvian paradigm and notions of proficiency and ownership in the outer circle. In this thesis, I argue that many educated individuals in the outer circle have inherited English in their societies (as represented by Kachru's model) and many have therefore appropriated English and are theoretically already proficient in English (communicatively competent) across a range of domains including education. This is explored in the first strand of the study and so I now discuss the concepts of proficiency and ownership as they relate to Kachru's model of 'World Englishes' (1985).

2.1.3 Proficiency, Ownership and Kachru's 'Three Circles'

Kachru's model has been criticised for its inability to define proficiency as part of the model of World Englishes, (Jenkins, 2009; Bruthiaux, 2003). However, the original purpose of the model was not to indicate proficiency levels but to illustrate the historical spread of the English language as well as the political status of the varieties. I now evaluate Kachru's model in response to criticisms but I also justify its use in this study in conceptualising a link between outer circle ownership of, and proficiency in English as well as the implications of this for academic language testing.

In explaining World Englishes, Kachru's model (1985) has been developed and critiqued since it was introduced, and indeed joined by other models (Galloway & Rose, 2015, p. 23). These include those of: McArthur (1987, p. 11); Gorlach (1988); Strevens (1992, p. 33); Modiano (1999, p. 25); Graddol (2006); and Schneider (2007). I use Kachru's original model

primarily as a useful way to conceptualise the historical and political legacy in the current ownership of, and proficiency in English of outer circle students. Essentially, using Kachru's model in my mixed methods study, I employ quantitative and qualitative methods to investigate the implications of *being* an outer circle student with regards demonstration of, and implications of prior proficiency in English for academia. In relation to the first strand of the research, Kachru's model is potentially useful in explaining test performance by nationality and the implications of this for admission to university because an assumption of needing to prove sufficient proficiency is based on nationality and qualifications. The model is also useful for conceptualising the target language use in the inner circle. The second strand of the research uses case studies to explore the ability of outer circle students to communicate and interact in this context and the relevance of their background in English in doing so.

Even though the model initially represented the historical "diasporas" of English-speaking and English-using communities across the globe as the English language spread (Kachru, 1985; Lowenberg, 1986; Pakir, 1997), the model can help to explain present-day ownership of and proficiency in English. As with others who appraise Kachru's paradigm while critiquing the actual model (Bruthiaux, 2003), I propose that Kachru's conceptualisation of the outer circle represents many who have been instructed and socialised in the medium of English in their multilingual, multicultural post-colonial states. Before I expand on this however, I must state that by using the model in my study, I am not suggesting that expanding circle countries do not possess any speakers and users of English who are as (or indeed more) familiar with and proficient in academic English as speakers from outer and inner circle countries. Neither am I suggesting that a more accurate picture of proficiency and ownership of English in the world today wouldn't be better expressed by another or new

model. Instead, I argue that a model is not intended to be accurate in illustrating exactly the linguistic reality on the ground, as Chapelle comments on "models" in reference to validity "since models are always simplifications of complex phenomena, the question is not whether they are accurate, but rather how wrong they have to be in order to not be useful" (Chapelle, 2012, p.27). Thus, my position is that it is still the case that the model carries with it a significance in relation to academic tests; in particular:

- Large numbers of students from outer circle countries are proficient in, and have an ownership of English, which is possibly represented in their English proficiency test scores (Bridgeman, Cho, & DiPietro, 2015; Zheng & Wei, 2014).

- Language proficiency tests tend to be based on inner circle norms - a practical choice in terms of test construction (Ackermann et al., 2010; Bamgbose, 1998; Shaw & Weir, 2007) but one which may be questioned on linguistic and political grounds (Davies, 2011a; Davies, Hamps-Lyons, & Kemp, 2003; Khan, 2009; Obaidul Hamid, 2014) .

- Alternative qualifications accepted by UK universities (appendix 3) tend to be from outer circle countries because of the historical legacy of English in former colonies (Read, 2015, p. 3).

It is important however, to first concede and articulate the flaws in Kachru's model. Kachru's model can be critiqued in terms of the variability of individual country characteristics and English use *within* and *between* the various countries and circles. By categorising countries into only one of three circles, his model has been deemed too "static" by some scholars (Taylor, 2009), especially due to the fact that is it based on geography rather than

sociolinguistic realities within countries (Bruthiaux, 2003; Jenkins, 2014). In critiquing Kachru's model, Jenkins (2014) claims that as the model was based on geography and genetics rather than identification with and use of English, there is a difficulty in using it as a way of defining the proficiency of speakers within and between those circles.

Regarding a lack of variability *between* countries in the model, it is true to say that not all outer circle nations display identical use of English in their societies and education systems. Tanzania for example followed a national policy of promoting Swahili in primary education and only used English as a medium of education in secondary and tertiary education. Although nominally Anglophone in terms of its association with fellow East African states Uganda and Kenya, Tanzania may actually lag behind in English proficiency as a result of Swahili being made the medium of education in primary schools in the 1960s and 1970s (Bwenge, 2012) and the country could be reclassified to expanding circle (Galloway & Rose, 2015, p.112). However, recently there has been evidence of a swing back to English (Ochieng, 2015). Another African country, Cameroon, is divided between an Anglophone and a Francophone population who could be said to straddle the outer and expanding circles. An Asian example mirrors the Tanzanian experience: Malaysia switched from English to Malay as a medium of instruction in national schools while at the same time maintaining a large number of private English-medium schools (Gill, 2005; Stephen, 2013).

It is also the case that outer circle countries can be divided into largely monolingual and multilingual societies. Relatively monolingual societies like Swaziland (Siswati), Lesotho (Sesotho), Botswana (Setswana), Bangladesh (Bengali), and Hong Kong (Cantonese) do have majority unifying languages (bracketed) other than English and may not always use English as a medium of instruction or indeed communication throughout their education systems and societies. There are also examples of countries such as Nepal that have not strictly been

considered outer circle, but where English does have a long history and is now considered by some as "indispensable" in the social and institutional life of the country (Giri, 2015). Other countries in the outer circle are highly multilingual such as Malaysia, Singapore, India, Zambia, Fiji, South Africa, Ghana, Kenya and Nigeria and do use English extensively in domains such as education systems and the media in the absence of a widely-accepted "native" unifying national language (although in Kenya Swahili does perform a unifying role to some extent).

Regarding a lack of variability *within* countries, Bruthiaux (2003) argues that Kachru's circles model has "outlived its usefulness" (ibid, p.161) because it does not differentiate between various English-speaking communities within outer and expanding circle countries. For example, there are many L1 English-speaking communities within the outer circle (often of south Asian origin) who have either inherited or adopted English as a native language in their ancestral lands and/or in their "new" homes via migration within the British Empire. Examples include English speakers in South Africa (Mesthrie, 1992), Anglo-Indians in India (Spencer, 1966), Eurasians and Indians in Singapore (Rubdy et al., 2008; Wee, 2002), and Goans, Ismailis, Parsees and others in Goa and East Africa (Kurzon, 2004; Salvadori, 1983). These are examples of the 'grey area' to which Jenkins refers (2009, p.20). Other examples include people of English-speaking Jamaican-descent in expanding circle Spanish-speaking central America (Holm, 1983). The post-imperial linguistic situation regarding English as a first language is clearly more complex than the inner circle/outer circle dichotomy would suggest and therefore the model's "over-simplification" of the sociolinguistic reality on the ground in outer circle countries is a fair criticism. Conversely, there are many communities in rural parts of the outer circle that could be defined as expanding circle with regards to the spread of English, to take Ekegusii speakers in rural Kenya as just one example (Michieka, 2009). Relating to this urban-rural divide, Kachru's paradigm has also been criticised as

being “elitist” through positioning outer circle varieties as “educated versions of local English” (Canagarajah, 1999, p. 180) so highlighting a discrepancy as to how local elites orient to their versions of English and how the mass of the population orients to English.

However, I would argue that the model is not supposed to differentiate between different users and uses of English within a country, and that Kachru’s later concept of “functional nativeness” (1998) would cover the range and depth of uses of English within a country (next section). Kachru actually acknowledges that the linguistic situation in the outer and expanding circles is constantly developing “The English-using speech communities are continuously growing, evolving, and changing” (Kachru, 1997, p. 69), and he always intended his model to capture the “historical, educational, and functional distinctiveness” of the use of English in different places (ibid, p.68).

Jenkins also critiqued the term “inner circle” which, she reasoned, implied that speakers in this circle are central to the concept of World Englishes whereas she counters that in fact their influence is on the decline (Jenkins, 2009, p. 21). However, regarding the inner circle as being “central” to the model, this is simply a reflection of present-day norms. The inner circle does still provide the norms for test construction (Ackermann et al, 2010) and the inner circle is still exempt from taking academic English proficiency tests because politically, they *are* still considered central to English language standards (at least by the UKVI). Whether this is right or wrong is another matter and has been discussed elsewhere (Hamps-Lyons & Davies, 2008; Khan, 2009; Obaidul Hamid, 2014). The division between the inner and outer circle is therefore political as well as potentially linguistic and still exists in the areas of my study that I focus on, particularly with regards university admissions and acceptance of standards and norms. Even though Kachru’s conceptualisation of World Englishes based on nationality has

been criticised, the historical dominance of the inner circle in language pedagogy and testing cannot be ignored or brushed aside. It exists in the numerous policies of governments (i.e. MESCs) and institutions and therefore there are issues worth exploring within this political reality such as patterns in test performance according to nationality.

Ultimately, Kachru's model does reflect historical realities in terms of the spread and inheritance of English in colonial societies and education systems, which has effects lasting to the present. I next look in detail at the outer circle to discuss how proficiency in the outer circle can be regarded.

2.1.4 The Kachru Paradigm – The 'Nativised' Outer Circle

Within the debates on world Englishes, there have been discussions and conceptualisations that suggest that we need to rethink who a native speaker is. There have been many voices arguing that the inner and outer circles possess a similar level of proficiency in English (Crystal, 2003, 2012; Davies, 2013; Graddol, 2006; Gupta, 2006; Higgins, 2003). Indeed, the range of scenarios and domains in which a speaker is actually using English may reveal to what degree a speaker of English is "proficient" in English. In answer to the criticisms by Jenkins and Bruthiaux above, Kachru had already written about "functional nativeness" (Kachru, 1998) in order to cover the question of proficiency between the circles and differentiate between and within varieties of English. This concept acknowledges that there is a difference between Kachru's actual three circles model representing "World Englishes" and his wider thinking on the subject (Chee, 2009).

Kachru's idea of functional nativeness (1998) states that users of English (from whatever circle) may be highly proficient in English according to the "range" and "depth" of their

English usage in their home society. Range refers to the domains, contexts or situations a person would use English in (and with whom) and depth refers to the extent to which they would use English within that context. For example, a speaker of English may have inherited an L1 that is not English, but if one looks at their actual use of English in their daily lives across a range of domains, then an individual may actually have functional nativeness in English.

Examples of range and depth as applied to the variety of English use within an outer circle country can be seen in a discussion of Indian English (D'Souza, 2001) where the author argues that English is used for a "wide range of purposes" (2001, p.146), and has "penetrated all layers of society" (ibid) even though it may only be used with "ease and fluency" by an elite. However, D'Souza also mentions the fact that many Indians now teach English abroad and that the language can be seen in many aspects of Indian life - humour, teenager talk, code-mixing, innovations in the language as well as lively discussions on what is "correct" or "incorrect" English. Thus she concludes that regarding English "Indians have a stake in it, lay claim to it, love it, hate it maybe, desire it and revel in their mastery over it. It is perhaps, along with Hindi, the only true Indian language" (2001, p.150). This is an articulation of the ownership of, and proficiency in English for Indian speakers.

Mirroring this concept of range and depth of use of a language, Leung et al (1997) refer to a speaker's "linguistic repertoire" in an orientation away from *native speaker* and *mother tongue* to a consideration of a speaker's knowledge of multiple languages "What is the learner's linguistic repertoire? Is the learner's relationship to these languages based on expertise, inheritance, affiliation, or a combination?" (p.555). In the case of the outer circle, the countries in question have "inherited" English from the former colonial power and in

many domains within those countries many individuals have expertise in, and affiliation with English.

Davies meanwhile (2013) developed the concept of a "native user" as a non-native speaker who only differs from a native speaker in the fact that his or her English has been acquired since childhood through education (ibid, p.3). This pattern of acquisition may account for cognitive differences with a native speaker but essentially Davies argues that through education, "non-native" and "native" speakers are not very different at all:

> A native user may be indistinguishable from a native speaker who is equivalent in terms of education and language proficiency, with the exception of accent perhaps, but who, because of the absence of what is often considered to be the defining characteristic of a native speaker, early acquisition from parents or guardians or carers, does not qualify as a native speaker (Davies, 2013 p.26)

Essentially, Davies argues that the native user can attain the same communicative competence as the native speaker through education and gain access or membership into a domain or native speaker group "through control of the standard language" (2013, p.viii). Standard language for Davies is isomorphic with the term "idealised native speaker" (2011, p.297-298) which is in many ways an unattainable ideal for both "native" and "non-native" speakers as both sets of speakers have their own idiolects and grammars. Davies acknowledges that the term 'native user' is "oxymoronic" (p.11) and he purposely binds together "native" (L1) with "user" (L2) to illustrate the type of individual for whom English is a second language. However, he seems to want to go beyond the meaning of English as a Second Language and suggests that the native user can catch up to and indeed exceed the native speaker in the use of English, *through education*. In this conceptualisation, Davies aligns with Hulstijn who states in relation to the CEFR that "the higher levels of most of the competence scales also will normally be attainable only by people with higher levels of

education or functioning in higher professions" (2011, p.240). This also implies that "native speakers" are not automatically proficient in academic English for university. As with Kachru's "functional nativeness", this term of "native user" can certainly describe outer circle *and* expanding circle speakers of English who reach a high proficiency in English despite not necessarily having had English as a first or equal bilingual language from childhood.

The idea of a merging of the native speaker and native user can also be found in Graddol's arguments about globalisation and English (2006) whereby he refers to the development of the Kachruvian model of the inner circle as now consisting of "highly proficient" speakers and users of English "regardless of how they learned or use the language" (p.110). Similarly, Modiano (1999) advocates "occupation" of the inner circle with "proficient speakers of English as an International Language" from wherever in the world (p.25). Other voices have questioned the dominance of the inner circle in defining who a native speaker is. Canagarajah (2006), for example talks about outer and expanding circle Englishes being "quite central to the character and currency of English today" in this era of "post-modern globalisation" (p.232). Reflecting the political-historical nature of Kachru's model, the term "native speaker" has also been called a political act so that nativeness is "geopolitical, not linguistic" (Freeman, 2015). Thus, the 'inner circle' is now less and less a fixed geographical concept than an international cross-border concept potentially including individuals from all the circles of English. This is due to the fact that many individuals in the outer *and* expanding circles are increasingly being socialised or educated in the medium of English so what is true for the outer circle would logically also be true for many individuals in the expanding circle. Individuals from any circle can have a high proficiency in English meaning it is problematic to define the inner circle as wholly "native/proficient" and outer and expanding circles as wholly "non-native/less proficient". Unlike the expanding circle however, in the case of the

outer circle, inheritance of English can be seen in the infrastructure of the state in education systems, national media, books and printing, culture and other areas which have affected how many educated individuals in the outer circle have acquired English. I now look at some historical and social characteristics relating to English proficiency in the outer circle as represented in Kachru's three circles model to further articulate the ownership of English in the outer circle.

2.1.5 New Nations, Social Class and Education in the Outer Circle

Migration within the Empire and the creation of "new" independent countries since the 1960s means that the modern post-colonial outer circle state is often characterised by highly multicultural and multilingual populations. Many newly independent outer circle states continued to use English as a lingua franca and medium of communication and education. Furthermore, sociolinguistic and demographic developments in many outer circle countries both pre- and post- independence, mean that Englishes, or new varieties of Englishes are L1s for significant numbers of speakers in outer circle countries. Many other speakers of English are "native users" through education and "functional natives" through socialisation, living and functioning in their multilingual states using English to communicate both internally and regionally (Kachru, 1998; Lowenberg, 1986).

In these new states, social class and education now often determine the choice of first languages. The elites and middle classes in many metropolitan areas of postcolonial states have adopted English, for example, the anglophone "Afro-Saxon" elites in East Africa (Mazrui, 2004). This is also true where the parents communicate in, and decide to bring up their children largely in English (Higgins, 2009). For example, this language shift occurs in the case of individuals raised in English but from different parental language groups in West

Africa (Igboanusi & Wolf, 2009; Offiong & Mensah, 2012) and South Africa (De Klerk & Bosch, 1998). However, there is a sense that English of a standard good enough for university study is still very much the preserve of an urban middle class and elite in Commonwealth countries such as India and Nigeria "English, however, is an increasingly urban language, associated with growing middle classes, metropolitan workplaces and city lifestyles" (Graddol, 2006 p.50). English is connected with education and upward mobility in many outer circle nations such as Kenya "Asking people if they know English is almost synonymous with asking them if they have been to school" (Michieka, 2005, p. 179) and Malawi "For such Malawians, English is almost equal to education" (Kamwendo, 2003, p. 33). This pattern is particularly common across Africa where there are emerging and expanding middle classes (Kamwangamalu & Tovares, 2016; Matiki, 2001; Michieka, 2005; Ochieng, 2015). In socio-economic terms it is indeed the wealthy middle classes in outer circle and increasingly expanding circle countries who can afford to send their children to English-medium schools in their countries and eventually universities in inner circle countries.

There is thus, a metropolitan, internationally mobile class of people who are already proficient in English and function in English in a variety of domains including academia. Kachru's model therefore could be said to represent the continued significance of English in Anglophone Africa and parts of Anglophone Asia where affluent middle classes and elites in multilingual and multicultural states use English as a de jure or de facto national language (Graddol, 2006). Furthermore, as a legacy of Empire, countries in the category of the outer circle are more likely to have (or have had) school systems that teach in the medium of English inherited from Britain (including examinations):

> The newly independent Commonwealth nations generally inherited from the colonial authorities an English-medium education system, at least in the form of secondary schools and tertiary institutions for the political and economic elites. (Read, 2015 p.3)

In theory these systems maintained the school qualifications as *equivalent* to GCSEs or A-levels especially as they were often inherited and developed from examination boards such as UCLES or the University of London as in the case of Nigeria and their West African Examination Council (WAEC):

> The ordinances charged the council with determining the examinations required in the public interest in West Africa, and empowered it to conduct such examinations and to award certificates, provided that these certificates did not represent a lower standard of attainment than equivalent certificates of examination authorities in the United Kingdom (WAEC, 2000, p. 14)

This inheritance of qualifications and systems is reflected in the current use of Commonwealth CSELTs in the admissions process to UK universities (mentioned previously). For instance, when a university wants to ascertain the English language proficiency of students from the outer circle or Commonwealth, a whole range of Country-specific English Language Tests (CSELTs) are still referred to such as the Malaysian SPM; Nigerian WAEC; or the Kenyan KCSE (appendix 3). In the past, when students from the Commonwealth came to the UK for Higher Education they would usually be accepted for studies based on these qualifications but the variability in proficiency in the outer circle has led to the current situation whereby IAETs tests are argued as still needed (Read, 2015, p.3). One of the reasons for not waiving the requirement for outer circle students to sit IAETs is that students from outer circle countries can possess the full range of English-language ability - that of a native speaker, native user, EFL learner or non-speaker of English. There may be justification therefore, for arguing that proficiency in English through education or socialisation is not a truism for all students from all Commonwealth countries (Howell et al., 2012, p. 33). This is evident in the sense that most of the population in outer circle countries includes both urban and rural poor who do not necessarily have access to good quality

education as in Zambia and Malawi to take only a couple of examples (Williams, 2011, p. 44).

However, it cannot be ignored that many of the individuals from the outer circle who come to the UK to pursue higher education tend to be middle class and have already had experience of schooling in the medium of English, as well as use of English as an L1. Since independence, their nations have not rejected but appropriated the language, especially through education. Further examples of this phenomenon from the outer circle include Botswana (Alimi, 2011) and Brunei (O'Hara-Davies, 2010). All of the historical and social factors discussed above mean that in the outer circle, the boundary between native and non-native and first, second and third language speakers has become blurred (Crystal, 2012, p. 177). This blurring can be expressed in the concepts of "nativised" varieties of Englishes leading to Kachru's "functional native" (1998) Davies's "native user" (2013) and Graddol's "highly proficient speaker" (2006). This proficiency is part of the overall concept of 'ownership' or appropriation of English by people in the outer circle (Norton, 1997, Widdowson, 1994, Higgins, 2003) and I deal with this aspect in the next section.

2.1.6 Ownership and Language

Higgins makes the case for ownership being a better way to describe English speakers' proficiency because it avoids the "overly static dichotomies that inner-outer circle, or NS-NNS, produce" (2003, p.619). Another way of thinking about proficiency and ownership is that of a "legitimate speaker" (Norton, 1997) of a language. Norton (1997) cites Bourdieu's concept of a speaker of English or any language for that matter having "legitimacy" or of being a "legitimate speaker" (Bourdieu, 1977). Norton states: "If learners of English cannot claim ownership of a language, they might not consider themselves legitimate speakers

(Bourdieu, 1977) of that language" (1997, p.422). In her related thinking on "investment" in a language with regards to social identity, Norton frames a question: "How is the learner's relationship to the target language socially and historically constructed?" (Norton 1997, p.411). These questions of legitimacy and investment could be answered for the outer circle in that many individuals in outer circle societies have inherited English through their national identities and personal experiences in the domestic, public and educational domains of use of English.

For Seilhamer (2015) and his investigation of the ownership of English in an expanding circle country, Taiwan, he applies a framework of ownership of prevalent usage, affective belonging, and legitimate knowledge as a way of defining ownership - "usage" and "knowledge" also pertain to proficiency. He found that his cases displayed varied levels of ownership which also related to their socio-economic status. He also refers to the much more noticeable ownership of English in Singapore, contradicting the official government line that there are no native speakers of English in Singapore - cited in a study by Wee (2002) - by countering that:

> Widespread Singaporean usage of English, an emotional attachment to English by a substantial portion of Singaporeans, and an increasing sense of linguistic authority, especially among younger Singaporeans who have learned the language from infancy, all collectively suggest that English may be in the process of becoming a truly local language in Singapore. (Seilhamer, 2015, p.373)

This is an articulation of ownership for Singaporean English speakers. But for others, "authority" over the language may be a stumbling block to ownership. This is because Seilhamer makes the case that for many post-colonial societies such as Malaysia, individuals still look to the UK for authority on the language and do not therefore say that English is "theirs" if we consider authority (on linguistic accuracy or "correctness") as part of owning a language. Yet for Seilhamer, "ownership" remains an elusive concept in need of a "coherent

framework" and is an important and growing concept. Individuals in his studies are often located somewhere along an imagined scale of ownership with no ownership and ownership at each end:

> Between these two extremes lie feelings of comfortable familiarity with a language, assertions of legitimate speaker-hood, and instances of playful manipulation of the language, even when individuals stop short of making overt ownership claims. (2015, p.375)

One aspect of this growing concept is that ownership of English may differ according to place, society, colonial experience of English and past and current language policies. The majority of Malays might not claim possession of English due to their high levels of literacy in Malay and the official switch from English to Malay in the education system since the 1970s. However, it is still the case that English often plays a very different role in many outer circle African countries where it is often the language of national unity and literacy in highly multilingual nations where proficiency in mother tongues is often oral but not written - despite best intentions to promote literacy in African languages (British Council, 2013).

Regarding the link between proficiency in a language and ownership, Widdowson says:

> "Real proficiency is when you are able to take possession of the language, turn it to your advantage, and make it real for you. This is what mastery means. So in a way, proficiency only comes with nonconformity, when you can take the initiative and strike out on your own" (1994, p.384)

He later suggests in a talk given in Senegal, West Africa that outer circle nations such as Nigeria and Ghana "may well wish to appropriate the language" (1997, p.141). There is a suggestion here that this process makes English a highly distinct localised variety – what Higgins describes as "ownership as indigenization" (2003, p.619) but Widdowson also states that where English is adopted in countries "it does not follow logically, however, that the language will disperse into mutually unintelligible varieties" (1994, p.385).

Gupta (2006) asserts that inner and outer circle Englishes have become "functionally and attitudinally similar" stressing the ownership of English "English belongs to its people in the Outer Circle, just as much as to its speakers in the Inner Circle" (p.95). This link between proficiency and ownership is an important aspect of English in countries where individuals have appropriated English despite an often difficult colonial inheritance (Alimi, 2011, O'Hara-Davies, 2010, Higgins, 2009). Outer circle students are also proficient in different varieties of English, whether a localised one (Pidgin English in West Africa, "Singlish" in Singapore) or a more standardised and internationally intelligible outer circle form (Gupta, 2006, Rubdy et al, 2008) and it has been suggested by Canagarajah that proficiency is an ability to negotiate other varieties of English as well as being proficient in your "home" variety - an "ability to shuttle between different varieties of English and different speech communities" (2006, p.233). This is something that many multilingual individuals in the outer circle do on a daily basis.

Higgins (2003) makes the extension that outer circle speakers of English are not the only ones who could claim ownership of English "Even expanding circle speakers from nations such as Korea or Brazil may have high degrees of ownership, particularly those who are educated in private, English-medium schools or those whose socioeconomic status affords them ownership of English" (p.641). Ownership can therefore cover any individual who has had experiences and a background relationship with English (albeit often a privileged one) from British post-colonial societies to the expanding circle "with regard to English, it is no longer solely *their* language, but belongs to all those who use it in the four corners of the world" (Schmitz, 2012, pp. 251-253). I follow the stance of researchers such as Norton, Higgins, Seilhamer and Widdowson arguing that ownership is tangible in the societies of many outer circle post-colonial nation states. I include this ownership or appropriation of

English to mean proficiency in both a localised or colloquial variety and a more standardized, internationally-intelligible national variety, as a reality for many outer circle students, particularly with regards to multilingual African states as well as Asian states such as India, Malaysia and Singapore. In my conceptualisation of ownership of English in the outer circle, ownership is not just attitudinal but is also bound up with the history of the spread of English, inherited in a national as well as an individual sense, and is also concerned with socialisation in English, in education as well as in the home. Given that many outer circle individuals are proficient in English to the extent of being "native users" (Davies, 2013) as well as native speakers, they have ownership of English through being "legitimate speakers" (Norton, 1997; Higgins, 2003) of the language. In this way, I have articulated that the notion of ownership is to be thought of as inseparably part of the proficiency of the outer circle "native user" or "native speaker" of English.

2.1.7 Implications of Proficiency and Ownership for Language Tests and Admissions

This "picture" of proficiency in, and ownership of English in the outer circle stemming from the inheritance of English in a range of domains and aspects of life poses challenges for tests and university admissions in terms of moving away from categorisation of individuals as either native speakers (NS) or non native speakers (NNS) of English. OC test takers may be either native speakers; possess a "functional nativeness" in English (Kachru, 1998); be native users (Davies, 2013) (highly proficient ESL speakers); or they may well be EFL learners and speakers. Students from the outer circle are therefore often caught in a conceptual "no man's land" between a native speaker and a non native speaker whereby their variety of English is legitimate to them and in their country, but not to the philosophies behind the design and application of academic English tests and admissions policies. Ownership of a language

poses interesting questions for language tests and university admissions including how outer circle students perform on academic English language tests and how the results can be interpreted; what outer circle students feel about taking these types of test and how their performance on tests compares to expanding circle students.

There are signs that many outer circle students are uneasy with an increasingly ambiguous situation of having to demonstrate proof of proficiency despite evidence pointing to their existing ownership and proficiency. The UK Council for International Student Affairs, in a 2011 report stated that "Many students from countries such as Singapore and South Africa were affronted at being required to take a test having lived in majority English speaking countries, having done all their education in English" (ibid, p.16-17). This "affronted" feeling is best summed up by anonymous student quotes in the UKCISA report:

> I have completed all my studies since the age of 7 in the English medium and yet I had to take an IELTS test & on arrival in the UK the Immigration Officer did not speak English to my level!!!!!!!
>
> I am a native speaker of English, but from SA. If I hadn't already done an MA I would've had to do an IELTS test. For which I was a certified examiner for three years or so. Completely ridiculous.
>
> (UK Council for International Student Affairs, 2011, p. 17)

UKCISA concluded that "The system of SELTs is insufficiently flexible to take account of these cases in the way that institutional admissions systems routinely do, and creates an additional burden on students" (UKCISA, 2011, p.16). There was a recommendation for flexibility in admissions in point 9 of the same report to "Consider how the system can more flexibly recognise native English speakers and those whose education has been wholly or substantially in English" (UKCISA, 2011, p.5). Previously, individual universities did "flexibly recognise" individuals but the introduction of SELTs has in part led to the experiences and predicament of the students quoted above. It is important then to investigate

the proficiency of outer circle students as interpreted from performance on academic tests. For some of them it may be a valid way to test their English for academia while for others the taking of an academic test might be considered an unfair political and bureaucratic exercise.

2.1.8 Concluding Remarks

The term outer circle is an apt description in language testing because in the inner circle, outer circle varieties of Englishes are indeed left outside the norms. There is then, the incongruous situation whereby some outer circle students, despite their "functional nativeness" (Kachru, 1998) and backgrounds in English, are in the position of having to take academic English tests to demonstrate that they are ready to perform in an English-medium environment. It is incongruous because many individuals from the outer circle will already have been educated and socialised in English-medium education systems.

However, in this study it is important to look at these individuals' test performance and academic performance to see if the tests actually do highlight issues in their language proficiency regarding inner circle norms in academia. The academic English tests are positioned to be valid in this regard and are simply providing a tool for policy makers. There may still be a need to measure the English of speakers in the outer circle regarding their proficiency for inner circle academia, and the way that is done is by using international tests, which are based on *inner* circle norms. The question is then, what do these tests "say" about the proficiency of outer circle students, and do the tests have predictive validity for this outer circle demographic of student? The specific gap and research questions for this strand will be presented at the end of this chapter after a discussion of the second strand.

Regardless of where the student is from, and regardless of whether English is his or her L1, they will have to possess the right linguistic and academic skills in order to negotiate their respective courses at university. I now analyse how validity in testing and performance in academia is conceptualised.

2.2 Research Strand 2 – Predictive Validity and Academic Performance

The purpose of this section is to outline the concepts and literature behind the second strand of investigation of the study: an exploration of the predictive validity of the PTEA as it relates to individual students' academic performance in an inner circle context (the domain). In order to do this, I first discuss what validity is in the context of university admissions tests including discussion of predictive validity as part of criterion validity, authenticity and considerations of the characteristics of the test taker. This, then, leads to a discussion of how I conceptualise performance in academia before I review a variety of studies which have explored the concepts of predictive validity and performance in academia in past studies. This review allows investigation of the concepts, methods, questions and findings from previous studies that inform this study.

Predictive validity is concerned with interpreting test scores and comparing this interpretation to evidence of performance in the post-test context. It is concerned with making inferences and making an attempt to "predict" how someone will perform in the future domain. Before outlining this approach I introduce and discuss validity, a core concept of language testing and assessment, and its relevance to this study.

2.2.1 Validity

The focus of this study is on predictive validity which was borne out of the bigger concept of validity. In language testing, the concept of validity is the study of ascertaining whether tests do what they claim to do. Validation is acknowledged as being complex but "the basic questions are straightforward: First, what is being claimed? Second, are these claims warranted, given all of the evidence?" (Kane, 2011, p. 4). Validation is the process of seeking evidence to support the *inferences* we make about an individual's ability based on test scores. For example, in the case of proficiency tests such as the PTEA, the claim is that scores aligned to B2 can be interpreted to mean that a student will not be disadvantaged due to a deficit in language. In this way, score profiles are interpreted by the test takers themselves, admissions staff, teachers and other stakeholders to infer whether a student has the language proficiency to be able to cope on their academic course.

Students are admitted to courses based on a minimal attainment of a test score aligned to B2 in the CEFR (allowing individuals to infer – rightly or wrongly - that students will not be disadvantaged on their courses because of language). The predictive validity evidence collected to support these inferences involves looking at student performance on their courses and comparing this with their academic English test scores. Before discussing predictive validity and evidence collection, I first outline the origins of validity from the field of psychology which produced the concept of criterion emphasis which led to predictive validity.

2.2.1.1 Criterion and inferences

Cronbach worked in the arena of psychological testing and he developed validity in reference to how a test is interpreted and so validity is about interpretations or *inferences* made from a test. Interpreting scores and making inferences from scores are crucial in predictive validity. Users of a test must ask whether the inferences they are making from the test scores have validity. This depends on the decision to be made as "a test relevant to one decision may have no value for another" (1990, p. 150). Cronbach's conceptualisation of validity or "validation" was first initiated in his ideas surrounding logical validity and empirical validity in the first edition of his *Essentials of Psychological Testing* published in 1949. This was then developed into four aspects (content, predictive, concurrent and construct) with other researchers on the American Psychological Association (APA) Committee on Test Standards (1954). This was reduced to three aspects in his other work with Meehl (1955) and his final three emphases in validity in the 5th edition of his book (1990). Cronbach came to the conclusion that the process of validation is "an *inquiry* into the soundness of the interpretations proposed for scores from a test" (1990, p.145), an inquiry comprising a consideration of three aspects or emphases:

- Criterion emphasis (made up of predictive emphasis and concurrent emphasis)
- Content emphasis
- Construct emphasis

These terms have become important in subsequent discussions of validity and as such I outline them here.

The first emphasis, criterion, includes the two elements of *predictive* and *concurrent* emphasis in validity. The use of the term criterion refers to "an external variable such as a syllabus, teacher's judgement, performance in the real world, or another test" (Davies et al, 1999, p.34-35). Use of a criterion is about proving some sort of correlation between test scores (first criterion) and the second criterion (other evidence of proficiency), and was originally positivist in approach. However, in current validity studies the criterion can include qualitative elements such as opinions and perspectives, depending on how "performance in the real world" (ibid) is conceptualised and in this way can now be interpretivist. Both elements of predictive and concurrent criterion validity involve collecting evidence to support the inferences of a test. The dual nature of Cronbach's criterion validity is presented in his 1955 paper with Meehl:

> "The investigator is primarily interested in some criterion which he wishes to predict. He administers the test, obtains an independent criterion measure on the same subjects, and computes a correlation. If the criterion is obtained some time after the test is given, he is studying *predictive validity.* If the test score and criterion score are determined at essentially the same time, he is studying *concurrent validity*" (Cronbach & Meehl, 1955, p.282)

With reference to my study, the criterion I wish to predict is academic performance, and the role played in that performance by a student's English language proficiency. I am inferring from test scores whether students will be able to cope with the language demands of their course. The independent criterion measure is the assessment of the student's proficiency in the domain (in comparison with the test score). I collect the evidence of proficiency in the domain "some time after the test is given" (ibid, p.282).

To illustrate his concept of predictive validity, Cronbach (1990), used the example of arithmetic aptitude tests giving a score from 1-50 compared with later school grades from A-D. His conceptualisation of predictive validity was the probability of getting a grade A-D

from a particular band in the aptitude test. A test score indicates the likelihood of getting a grade from A-D following on from the experiences of similar test takers who "previously scored at that level" (ibid, p.152). Therefore, in a predictive validity study, the test scores compared against the actual grades of participants can be used to anticipate what might be the case for others who score at that level. However, Cronbach talked about a second dimension to predictive validity in what he terms a "follow-up study" (ibid, p.153). He states that the validity of predictions is "checked" by a follow-up which "compares test scores with another measure" (p.153) which is a criterion deemed important by the "audience". This measure could include use of other criteria than that of grades. For example the "audience" or stakeholders might be interested in how students perform on specific tasks in the domain from the perspective of tutors as the criterion.

The second part of the criterion emphasis, concurrent, is concerned with whether a test can be compared with another construct such as another test assessing the same abilities or a syllabus as with the definition above (Davies et al, 1999). In validity studies of academic proficiency tests, this has been demonstrated in comparisons of test scores from different academic proficiency tests (Educational Testing Service, 2010; Riazi, 2013).

The second emphasis, content, relates to comparison of the test content with domain duties or tasks - whether the content of the test mirrors or adequately reflects the actual tasks that the test taker will be doing in the post-test situation, whether the test *replicates* or *simulates* in some way the tasks and conditions of the future domain.

The third emphasis, construct, is concerned with the *processes* behind the content – the "concepts" tested for. For example, in the academic domain students need to be able to

display *critical thinking skills*, and in this sense it needs to be asked whether this process or construct is present in the actual ability of students to complete the tasks in the actual test.

For Cronbach, the four types of validity above (criterion – predictive; criterion – concurrent; content; and construct) are differentiated in terms of the *emphasis*. In predictive validity, the emphasis is on the criterion *post-test*, with less concern for the content of the test (content validity) even though the two are connected. The idea of inferential validity emerges from Cronbach's writing. This is the idea that tests themselves can never be valid, but rather that it is the inferences we make about test scores that can be valid or invalid. To infer is "to deduce from evidence and reasoning rather than from explicit statements" (OUP, 2002, p. 725), which means that a score does not explicitly state someone is proficient or non-proficient, instead, evidence is collected (criterion evidence) as to whether the inferences we make based on an individual's scores actually support what we are saying about their proficiency. In this regard, predictive validity is concerned with "inferences about the likelihood of success in particular settings in the future" (Linn, 1990, p. 301). In this study I focus on the performance of individuals in the setting of university. Kane (2011) stresses that the "proposed interpretations" of test scores be clearly stated so that the claims and decisions based on test results are validated and not the test itself.

2.2.1.2 A 'unified' Concept of Validity and Social Consequences

Messick continued themes from Cronbach, particularly the idea of inferences but also developed a "unified concept of validity" (1989). He developed his own definition of validity, which continues to be influential:

> Validity is an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the *adequacy* and *appropriateness* of *inferences* and *actions* based on test scores or other modes of assessment. (1989, p.13)

For Messick, the process of validation is considered to be "scientific" in that inferences are compared to hypotheses and can be tested. In this way, as referred to above, empirical evidence should support the inferences. This evidence refers to the three types of validity essentially the same as that articulated by Cronbach. However, Messick's unified concept of validity is centered on construct validity embracing "all forms of validity evidence" (1989, p.17) because it incorporates content relevance coupled with inferences made about the future domain – whether the tasks can simulate in some way the future domain. It can be seen that this is about prediction.

The key issues in validity for Messick are "the interpretability, relevance and utility of scores" (p.13). Because of this unified view of validity, Messick's views on validation have been critiqued as being more philosophical than being able to provide practical solutions (Davies, 2011b) and in this way validity as an idea or all-encompassing concept is differentiated from the actual process of "validating" a test (validation). In Weir's "socio-cognitive framework" (2005) validity is conceptualised as incorporating context and criterion-referenced validity. Context validity is used interchangeably with content validity but with the added dimension that context "better accounts for the social dimension of language use" (ibid, p.19). Weir's definition of criterion-referenced validity mirrors that of Messick (1989, p.16) in including predictive and concurrent validity. However Weir talks about predictive validity as the "predictive power" of a test and is sceptical as to the ability of predictive validity studies to be effective mainly because of the practical problems associated with collecting evidence of academic performance and the abundance of other variables that affect academic performance (Weir, 2005, p.36). Despite this conclusion, he has since

undertaken a predictive validity study looking at criterion-validity and "real life academic performance" (2013), the results of which are reviewed below.

2.2.1.3 Test Use, and Predictive Validity

Bachman (2004) largely supports Messick (1989) in arguing for a "unified" concept of validity but with a stress on test use "validity is a single quality of the ways in which we use a particular test" (2004, p.260). Furthermore, incorporating Messick's key concept of the use of test scores, Bachman developed the idea of inferential validity relating to Messick's discussions on how a test is used and interpreted. In so doing, Bachman extends the discussion of predictive validity (calling it predictive utility) as part of his discussion of "authenticity" and the "Real Life approach" to language testing (1990), and later his "test usefulness" or "test use" (2005; 1996) and "Assessment Use Argument (AUA)" (Wang, Choi, Schimdgall, & Bachman, 2012) which emphasise collection of evidence that justifies a particular use of a test. Fulcher & Davidson (2007) put forward a similar idea of "pragmatic validity" (2007, p.18) whereby there is no absolute answer to the question of a validity of a test. Validity depends on the evidence that supports the use of a test, and the arguments made about what the evidence suggests about the test. A language tester collects evidence to support the use and interpretation of a test for an audience, however this use and interpretation may change as the evidence changes (the pragmatic aspect).

Chapelle claims that a definition of validity as a test measuring "what it purports to measure" is at odds with current definitions and in fact a better way to state the purpose of validity is to "examine how test users know what the score means and how it can be used" (2012, p. 21). This is a definition that also centres on interpretation of the test score as the focus of validity.

As such, Chapelle asks how this may be done, because a conceptualisation of "valid" being the equivalent of "good" is too simplistic. Instead, it is up to the researcher to "explore their own knowledge and beliefs about score meaning" (ibid, p.26), and presumably to argue for a particular interpretation of a score. Chapelle argues that validity is a constructivist concept because it relies upon interpretation of scores and is dependent on arguments and presentation of evidence. In this way, Chapelle is also a pragmatist towards validity.

2.2.1.4 Authenticity

Bachman makes the connection between the concepts of authenticity and validity and calls them "virtually identical" (1990, p.307). The assumption here is that if a test is authentic then it should reflect in some way the construct of the target domain (real life) and should therefore have some sort of predictive validity (or utility) based on how we use the test scores. Bachman's concepts are important to this study because his concept of validity unites the test content and the future domain in considering the predictive validity of the test. I expand below on the idea of authenticity for this study as it is important in defining the future domain from which I collect evidence of proficiency in this strand of the research.

Bachman calls the "real-life approach" (1990, p.301) to defining authenticity as "the extent to which test performance replicates some specified non-test language performance" (Bachman, 1990, p. 301). He connects this with "the accuracy with which test performance predicts future non-test performance" (ibid, p.302). According to Davies et al this real life approach to authenticity is:

> "An approach to defining language proficiency in which no definition of language proficiency is attempted. Instead, a domain of real-life language use which is considered to be characteristic of the performance of competent language users is identified." (Davies et al, 1999 p.164)

This connecting of the ideas of real life, authenticity and performance is important because a predictive validity study seeks to compare test scores with a measure of academic performance. Therefore, there is a need to visualise what performance in "real life" looks like and how it can be measured. This involves outlining the academic domain, articulating what tasks and functions someone typically carries out with the language, in various situations and with whom. In academia, the domain is university but this includes various sub-domains in lecture theatres, seminars, tutorials, social situations and the many interactions between the various individuals in that context.

Language tests are deemed authentic and therefore reflective of "real life" and "valid" if their content and construct match that of the domain or context in which the test taker is likely to find themselves in the future, post-test. By being "authentic" as part of their validity claims, test developers strive to demonstrate to stakeholders that the content and construct in the tests have some "real world" applicability or relevance to the future domain of academia. In this way, authenticity is incorporated into the validity and predictive validity of a test. As a predictive validity-criterion study, the focus of this study is on the interpretation of test scores and their relation to a student's performance in the domain and not the content of the test itself. For the purposes of this study, I am not investigating the correspondence between the tasks in the test and the future domain but am interested in defining the "real life" of the future domain as part of collecting evidence for predictive validity.

2.2.1.5 Authenticity, Outer Circle Englishes and Predictive Validity

An argument could be made for a conceptualisation of authenticity that includes the outer and expanding circle Englishes as a legitimate part of the domain of university or academia. For

example, In the Pearson 'Relevant' Fact sheet (2012e), Pearson claim that the PTEA reflects the "real life demands of study", largely through three aspects:

1. Using integrated skills
2. Using 'genuine' academic materials
3. Using 'International' academic English

Here, the test is portrayed as a reflection of the domain of academia in the inner circle not only in the use of the PTEA corpus (Ackermann et al, 2010) but also in the use of "international" English. This is reflected in the test content with the test claiming to be highly relevant for international test-takers because it incorporates 'non-native' English *accents* reflecting modern-day international universities (Pearson, 2012e). Pearson make an inclusion of non-native Englishes in terms of accents in their test content, a specific attempt at being modern and global "The use of international and non-native English makes PTE Academic a more appropriate test for modern global universities of today" (ibid, 2012e, p.3). However, Pearson reason that the fact that the PTEA has included English corpora only from inner circle countries is enough to give the test an "international flavour" (ibid, p.3). Pearson state that "using the main varieties of English ensures that all test takers are on an equal footing" and that "tests that use only one specific English variety put test takers not familiar with that variety at a clear disadvantage" (ibid, p.3). However, inclusion of "varieties" does not extend to outer circle varieties of English, so if a student from India, Kenya or Nigeria has been accustomed to their local varieties of English during their schooling or undergraduate degrees, these will not be reflected in the corpus of Englishes used in the PTE Academic. The inclusion of diverse accents found in the domain is however a concession to World Englishes:

> Furthermore PTE Academic is the only academic English test to include non-native English accents. This reflects the diversity of English that students are likely to experience at any university where English is the medium of instruction, and where students will be taught by professors and teaching assistants who are not native speakers of English. This use of international varieties of English and non-native accents demonstrates that PTE Academic is highly relevant to today's modern international academic institutions (Pearson, 2012e, p. 3)

The inclusion of international accents is perhaps an attempt to be authentic and to improve the perceived validity of the test among test users – the "face validity" of the test (Bachman, 1990, p.301-302). However, the statement above also sets out the target domain as not just a domain in which the student will encounter native or inner circle varieties of English, but also those of outer and expanding circle countries.

The concept of authenticity applied to outer circle Englishes can therefore include their acceptance in academic English tests for defining the domain. However in constructing academic English tests, content is usually based on inner circle varieties (Taylor, 2006, 2009). This is mainly because only inner circle varieties or norms (Davies et al., 2003) of English have been standardised and codified in contrast to those of the outer circle which have not (Pakir, 1997, p. 174). If a language has been standardised and codified then there is recognition of agreed standards for use in domains such as education. This has implications for the pragmatics of creating tests. Outer circle varieties currently have "no global currency" (Widdowson, 1997, p. 144), or "cultural capital" (Bourdieu, 1973), no prestige outside (or indeed often inside) of their own country, and so are not considered legitimate varieties of English for inclusion in the inner circle content of language tests.

In contributing towards the debate on consequences and values in language testing, Chapelle also acknowledges the problems inherent in the native speaker norm in testing and the selection of "privileged varieties" as standard (2012, p.29) – a reference to the inner circle/outer circle divide explored in this study. This leads to discussion of one more facet of

validity at the end of this section, namely, who the target demographic of the test is in terms of their test taker characteristics.

Test producers themselves are caught in a dilemma between making sure there are standard varieties which have been codified on the one hand, (Taylor, 2009) whilst maintaining respect for non-codified World Englishes on the other (Davies et al., 2003; Obaidul Hamid, 2014) but there has been some attempt to incorporate aspects of World Englishes in order to make tests authentic and reflective of the real world at university (Pearson, 2012e). However, practical considerations are important when creating tests and so it is difficult to create a valid test with an uncodified variety of English. Bamgbose sums up the situation for outer circle varieties "in the absence of a codified standard variety, examination bodies may play safe by avoiding any usage that may be considered by some people as non-standard" (Bamgbose, 1998, p. 4).

Furthermore, due to the sheer number of outer circle varieties, it is considered "impossible to take account" of every conceivable variety of English when designing language tests and so a limit has to be set on which Englishes to include (Shaw & Weir, 2007, p.17). Currently, the only codified and standardised varieties are from the inner circle However, as can be seen above in the stance of Pearson towards linguistic diversity in the test, tests do now need to consider non-technical issues in validity such as use of varieties of Englishes. This is due in large part to the work of Messick and Bachman in developing concepts around authenticity, and the social impacts and ethical concerns around test use. I now briefly look at these concepts in validity that are relevant to this study before discussing how aspects of validity and specifically predictive validity have been applied in studies.

2.2.1.6 The Socio-Ethical Dimension of Tests

Other theories on validity have focused on the social and ethical consequences and concerns. A major contribution of Messick to validity was an emphasis on the social values and social consequences of test use (1980) called by some "consequential validity" and in this way added to validity by stressing technical *and* ethical considerations in validity which others have since developed (Kunnan, 2000, 2010). By considering the consequences of using tests, Messick appeals to researchers to look at the social impacts tests have on individuals, and also on learning and teaching. Here, the concept includes washback – how teaching and learning may change due to the use (and misuse) of tests and test scores.

Kunnan's addition of "test fairness" to the social-ethical dimension of validity includes factors such as "race", gender, ethnic background and physical and mental disabilities in consideration of test design (construct), access and impact (Kunnan, 2000). Test fairness can also include consideration of varieties of Englishes in the sense that these varieties may need to be considered in the content of tests and that non-inclusion of outer circle varieties would "unfairly disadvantage" the outer circle test taker in international tests (Lowenberg, 2000, p. 43). I would extend Kunnan's conceptualisation and apply this interpretation of fairness in terms of admissions policies. For example there could be a challenge to a test on the grounds that students with certain linguistic backgrounds (outer circle varieties of English) are often obliged to take the test, whilst others with similar backgrounds (inner circle varieties of English) are exempt from the test. The background of test takers is an important consideration in validity in terms *who* is being asked to take the test and whether this is justified. I now discuss this aspect of validity in more detail as outer circle students are the focus of this study.

2.2.1.7 The Test Taker and Experiential Characteristics

The background and nature of the test taker in validity has been covered by some researchers (Hamps-Lyons & Davies, 2008, Bachman, 2004, Kunnan, 2000, Lowenberg, 2000). The concept of the test taker as a factor in validity includes possession of "traits" - psychological, and physical/physiological traits as well as "experiential" characteristics such as education, age, gender and first languages (O'Sullivan, 2000). I position the "outer circle" characteristic and medium of English factor as an experiential characteristic of test takers.

Bachman comments on test fairness and the nature of test takers by posing the question as to who should take language tests prior to an educational program and who should decide who should take them (Bachman, 2000, p. 40) but although he believes this characteristic of language test takers should be involved in this aspect of test fairness, he leaves it open for discussion. Later however, Bachman does conceptualise the experiential characteristics of test takers as "personal characteristics" (2004) including L1 (mother tongue) and gender as group characteristics. The nationality of the outer circle test taker is also deemed a personal characteristic which could be said to have a "systematic effect" on test scores (Bachman, 2004, p. 156) – expressing the idea that nationality may be a significant variable in test performance. This could suggest that the inferences made based on test scores could be "contaminated" due to "construct irrelevant variance" (Davies et al, 1999, p.32). Bachman appears to be saying that the variance of the scores on a test (such as nationality A displaying mean scores at C1 level and nationality B displaying mean scores at B1 level) could be explained by the background of the students (nationality) and not the construct of the test, which is academic English for non-native speakers of English. However, if this is so, then this may simply reflect the fact that test takers from one nationality do tend to have a higher proficiency in English than others. This can be investigated by looking at variation in test

scores according to nationality using statistical tests such as Analysis of Variance (ANOVA). In defence of tests, it might be argued that it is hard to cater for everyone's background in the construction and design of tests, particularly with the heterogeneous target audience of international tests (Shaw & Weir, 2007, p. 17). However, I suggest that because the test taker is from the outer circle with a corresponding inheritance of English and associated proficiency in English (ownership), these could then be significant factors in affecting their test scores in comparison to the ownership of, and related proficiency in English of test takers from the expanding circle.

2.2.1.8 Concluding Remarks

Test validity began as a very technical concept but now incorporates both technical and ethical considerations (Chapelle, 2012) reflected in the development of validity by various scholars since Cronbach. Technical considerations of content, construct and criterion (APA, 1954) have opened out into issues of social, consequential and ethical concerns: consequential validity, authenticity, test use and test fairness (Bachman & Palmer, 1996; Kunnan, 2000; Messick, 1980). Validity is now understood as related to the content of the test, its construction, assumptions and scores, how we interpret and use those scores, and in the evidence we collect to support arguments for the inferences we make about the test scores and uses. As an extension of the discussions around the use of the test, I include consideration of who the legitimate targets of the test are and consideration of their backgrounds in how they perform on the test.

Predictive validity can take into account these concerns but in order to ascertain predictive validity there needs to be a gathering of the language proficiency scores of students, followed by an attempt to make inferences based on those scores as to what the test-taker might

struggle with in the future, at university. The criterion of performance at university then needs to be defined in order to collect evidence for the inferences made on the test scores. This evidence takes the form of data from various sources. Inferences from the test scores themselves theoretically provide the evidence (if not always perfect) of proficiency, and the future domain must provide evidence of performance (criteria) with which to compare against the test scores that represent proficiency. This evidence can be both quantitative and qualitative in nature, and positivist and constructivist/interpretivist. As we will see later in the findings of past predictive validity studies, many studies have sought out quantitative evidence of performance but it is also possible and desirable to develop a mixed methods approach to determining predictive validity that incorporates qualitative evidence. In this study, the task is to provide evidence in support or otherwise, of the predictions based on what the students' score profiles suggest, e.g. that a student at B2 overall or in a particular score category will be able to cope on their course, and one at B1 level might struggle linguistically-speaking. I now define the criterion of "performance" in order to collect the predictive validity evidence.

2.2.2 Academic Performance

2.2.2.1 Aspects of Academic Performance

As articulated above, predictive validity is a type of criterion validity (APA, 1954; Cronbach, 1990; Messick, 1989; Weir, 2005). In predictive validity studies, ultimately, we want to collect evidence in order to know what we can *infer* from test scores, whether the scores 'say' anything about the student's abilities and how students will perform in the target domain.

The primary concept of academic performance in the context of this predictive validity study is linguistic performance in the university domain. That is, assessing whether a deficit in language proficiency (inferred from test scores) is detectable in the communication and interaction of the cases in the future domain. The concept of performance also considers whether any deficit in this language performance affects the proficiency encapsulated in academic socialisation, EAP "study skills" and academic literacy.

In investigating predictive validity we "extrapolate from universe scores to performance in domains beyond the test" (Bachman, 2004, p.263) and in doing so, the adequacy of extrapolating or making inferences from test scores depends on "multiple sources of empirical evidence" (Weideman, 2012, p. 6). In order for predictive validity to "work", a researcher collects evidence of linguistic and behavioural performance in the future or target domain of university "beyond" the test. It can then be asked whether the test scores (in their interpretability) marry against what can be observed in the test takers once they are in their future academic environment. A test provides scores that are interpreted to provide evidence that a test taker can handle the "criterion situation" (McNamara, 2006, p. 33) or "target language use situation" (Taylor, 2009) which is the domain of academia in this study.

Academic performance is a term that encompasses various elements. A student needs a certain level of general academic English proficiency to be able to communicate and to interact with people on any course of study. This is the type of proficiency tested for in academic English tests. However, in assessing a student's academic performance a student also needs a more general set of study skills needed for any subject and academic literacy for a particular subject (discussed above in proficiency).

Performance has been equated to "outcomes" (Andrade, 2006; Morrison, Merrick, Higgs, & Le Metais, 2005) that is usually defined as an attainment of grades at university. In predictive validity studies this is often measured in quantitative ways such as comparing language test scores to assignment grades, exams and test scores. In many studies "grade point averages" (GPA) and final academic grades have been used as *the* measure of performance. However, performance can also be measured in qualitative ways such as perceptions of performance in tasks or "study experiences" using the opinions and judgements of course tutors and the students themselves. In this section I first explore the ways in which studies have conceptualised academic performance in terms of quantitative measures. I then turn to the term "ability to cope" to deconstruct what I interpret the term to mean in relation to tasks and situations in academia before looking at the findings of past studies in predictive validity and academic performance.

2.2.2.2 Academic 'Outcomes' or 'Achievement'

Performance can be about academic achievement. This can be measured by gathering "evidence of learning, which may be measured by successful completion of course requirements, grade point average (GPA), satisfactory academic standing or retention" (Andrade, 2006, p.143). In determining academic performance related to achievement of grades, many studies have taken a quantitative approach using statistical analysis. These studies have traditionally sought a statistical relationship between the independent variable of language proficiency test scores and a dependent variable such as either a final academic grade, a grade point average, or an end-of-first semester grade. A grade point average or GPA is "an internationally recognised calculation used to find the average result of all grades achieved for your course" (Monash University, 2014). A quantitative measure of

performance often employs the use of statistics and *correlation.* A correlation is "an assessment of how strongly two pieces of data appear to be connected to the extent that a change in one of them must produce a change in the other" (Tapson, 2006, p. 116). This method typically sets out to find a *statistically significant* correlation between one set of scores – indicating linguistic proficiency - and another future set of scores – course grades and/or final grades. In past statistical studies of predictive validity studies, GPA or a final academic grade is often seen as *the* measure of academic success and this pattern for predictive validity has been outlined by Davies (2007):

> The normal method for assessing the predictive validity of these proficiency tests was by simple correlation (product moment) between the test (usually taken at the start of the academic year) and the degree or diploma examination taken at the end of the same academic year. (ibid, p.82)

This has been the main method for determining predictive validity in a range of studies in the 1980s, 1990s and 2000s. Descriptive statistics characterise the sample and then inferential statistics are used to calculate the statistical significance of any correlation found between an entry score and a final academic grade.

In terms of sample sizes for these statistical analyses, most studies have been modest, examples of these include; 101 undergraduates (UGs) and postgraduates (PGs) in Feast (2002); 82 PGs in Woodrow (2006); 65 UGs in Dooey & Oliver (2002); 62 1st year UGs and PGs in Hill et al (1999); 60 UGs and PGs in Riazi (2013). Others were larger; 171 UGs in Weir et al (2013); 196 UGs in Johnson (1988); 376 PGs in Light et al (1987). A couple of studies have had much larger samples; 6516 students in Wait & Gressel (2009) and 2594 UGs and PGs in Cho & Bridgeman (2012). Many of these studies have found a statistically significant, positive, but weak relationship between language proficiency test scores (overall

'global' scores or sub skill scores) and subsequent academic achievement in the form of final academic grades or grade point averages.

There are problems with correlation methods. A calculation of grade point averages as a comparison to the language test score does not take into account the number of subjects or modules studied or weighting of scores and marks on any particular course. Fox (2004) suggests that studies that only rely on a correlation of test scores and GPA are "particularly weak" because GPA is actually dependent on other factors such as "social networks of support, determination, financial security, time available for study, and time spent studying" (ibid, p.442) which introduces other variables into the cocktail of predictive validity. Other studies also cite many other factors beyond linguistic ability that affect performance. Ingram & Bayliss sum up the problem with predictive validity in their 2007 study:

> As IELTS measures only English language proficiency, attempts to correlate test results with subsequent academic results that depend on a multitude of other factors (intellectual ability, motivation, quality of teaching, learning style, acculturation, etc) will inevitably fail or, at best, be open to serious criticism (2007, p.5-6).

Wait & Gressel (2009) go so far as to say that a test is used beyond its scope if used for academic performance prediction and thus, it should only be used for proficiency scores for admittance purposes, not to predict academic performance. I would counter this view however, by observing that by the very act of using proficiency scores for admittance purposes, one is already making future inferences or assumptions of that individual's likelihood of success based on their test scores which need to be explored and justified.

Generally, what was revealed in the correlation statistical studies is that above a certain level of proficiency, language is less of a factor in academic performance or success and instead, other variables (such as those mentioned by Fox) become apparent "it seems likely that there

is a minimal level of English proficiency required before other factors assume more importance" (Graham, 1987:517). An aspect in common here is that many studies use a truncated sample of test takers. This refers to the fact that most sampling in predictive validity studies has only taken students currently *at* a university, therefore the sample is theoretically already of a high proficiency or standard in English. Test takers with lower proficiencies do not usually reach the research field and therefore it may be harder to make a link between linguistic proficiency and academic performance as all the lower levels have been "filtered out". This then causes a weak positive relationship in the studies' correlational analyses because the lower proficiency levels simply are not represented in the data – a problem cited by others (Cho & Bridgeman, 2012; Ferguson & White, 1992; Kerstjens & Nery, 2000). Cho & Bridgeman refer to this problem in predictive validity studies as "range restriction" (2012, p.425) and it can be concluded that a strong relationship between linguistic level and academic achievement may be hard to find as a result not least because "language may be more important in academic failure than success" (Ferguson & White, 1992:61). Because those with low proficiency scores are generally not at university and have no measure of academic achievement, it is usually only possible to examine samples at the threshold level of university (CEFR B1+ and B2 level) which is largely the focus of my study.

In other studies which did include lower proficiency students, stronger predictors of academic success were found among these students and the higher the language level, the more other factors apart from linguistic proficiency came into play (Cotton & Conrow, 1998; Ferguson & White, 1994; Graham, 1987; Elder, 1993; Kerstjens & Nery, 2000). A weak positive relationship between linguistic proficiency and academic grades is the norm in many of these studies because above a certain level of proficiency it is difficult to name language proficiency *alone* as a causative factor in academic success. If this were not so, every native

speaker would pass university degree courses with a distinction each time from the simple fact of being proficient in English. This is clearly not the case as evidenced in studies such as Dooey & Oliver (2002), Riazi (2013) and Lowe & Cook (2003).

Methodologically, there is a one-dimensional character to many predictive validity studies that only use quantitative methods to establish predictive validity. Banerjee (2003) suggests that traditional predictive validity studies are "out of date" or "bankrupt" in that they only look at a correlation between test scores and final academic grades or GPA. Banerjee suggests a widening of predictive validity in order to take account of the concepts of "study experiences" instead of only "study outcomes" (2003, p.45), and introduces the idea of the "cost" of being admitted to university with a low proficiency score (explained more below in 'actual experience'). In common with Banerjee, Ingram and Bayliss (2007) are more concerned about linking language test indications of proficiency to experiences in the form of "language behaviour" rather than a concept of academic performance.

A *purely* quantitative correlation study of academic success is limiting in its ability to explain the connection between linguistic proficiency and academic success because of the inability of this approach to account for all the factors that affect academic success. In order to be more effective at explaining relationships between proficiency and academic success, a predictive validity study needs to provide statistical analysis with a complementary approach using qualitative methods to explore the other variables involved in academic success. I will now outline what alternative concepts, methods and discoveries can be used to measure and document academic performance.

2.2.2.3 'Ability to Cope'

2.2.2.3.1 – Defining an ability to cope

An alternative way of exploring academic performance is to look at the "ability to cope" as a way of conceptualising the performance criterion. Academic language tests seek to measure whether test-takers have the right level of English for *coping* with university study in the target language (in this case, English). "To cope" means to "deal effectively with something difficult" (OUP, 2002, p.314) and in the context of performance in academia the phrase "ability to cope" is concerned with students' encountering of difficulties with tasks and situations and identifying whether this is due to a lack of sufficient language proficiency and/or a lack of appropriate skills. An ability to cope must be "measurable" and definable and should be "a central part of investigating validity" (Fulcher & Davidson, 2007, p.5).

An ability to cope with something indicates a readiness to be able to handle as yet unseen future difficulties. In order to be "ready" for academia, a student, whether an L1 speaker of English or not, not only needs the necessary language proficiency to be able to cope on their course but in addition, crucially, the academic skills and subject literacy with which to do so. It is therefore important to state that language proficiency aside, coping with university study is as much a challenge for a so-called native speaker as it is for a non-native speaker and can include factors such as a student's "background academic preparation in school, language proficiency, cultural expectations, and prior cognitive skills development" (Gilliver-Brown & Johnson, 2009, pp. 332-333).

Linguistically-speaking it is impossible to know beforehand the totality and range of activities and situations a student will find themselves in at university, but it is clear that a student does need a level of English which will give them the ability to cope, to "get by" in

the target language – crucially, *independently* – as outlined by the B2 CEFR descriptors (appendix 4). In this sense my study is mainly focused on whether the test taker has sufficient general English language proficiency. However, this soon "spills over" into discussions of academic literacy. For example, Fox (2004) undertook a study on an academic English test in Canada – the Canadian Academic English Language (CAEL) assessment. She states that "those who use the test draw inferences on the basis of CAEL test performances about a test-taker's ability to engage in and sustain academic work" (ibid, p.438). Regarding sufficient proficiency for academia, Fox acknowledges that there is a threshold "which is essential for engagement with academic work" (ibid) but says that the "quality" of the engagement is not guaranteed by the proficiency threshold.

Something more than a sufficient proficiency as inferred from the test is needed which suggests to me an ability to engage with a discipline and the skills needed in academia beyond a sufficient basic proficiency. I interpret the proficiency in the tests to be testing a general academic English, either "study skills" (Lea & Street, 1998), "generic skills and abilities" (Murray, 2010, p. 58), "basic interpersonal communication skills" (Cummins, 2008) or what Humphreys et al call a "general academic proficiency for pre-sessional students" (2012, p.35).

2.2.2.3.2 – Measuring an ability to cope

For the purposes of this study, there is a need to describe what an "ability to cope" may look like and an articulation of what a student should be able to cope with so that there can be some way of measuring it.

Fulcher & Davidson provide an example of ability to cope as a subject tutor's judgment of whether students can perform *tasks* such as "adequately read set texts to understand lectures and write assignments" (2007, p.5). A list of such tasks or abilities needed in order to perform at university using the target language can be seen on a variety of university pre-sessional course websites – these are courses that students undertake in language centres at university before entering university if their language proficiency levels are below that required by the university. Students on these courses (EAP) then work to improve their proficiency up to the required standard for university entrance before the term begins. These abilities or tasks on these EAP courses typically include:

- Reading widely and critically
- Summarising and communicating complex information in writing and speech
- Researching, planning, writing and editing assignments
- Referencing and citing accurately
- Presenting work, and demonstrating linguistic confidence in spoken production.
- Using information technology
- Using language appropriate to the situation and audience

The criterion situation or 'target language use situation' can be conceptualised in terms of the numerous situations and tasks that students find themselves in at university. This gives researchers ample contexts in which to collect data to provide evidence for "performance" in the domain. Looking through the literature of studies that have collected data on student performance, situations and tasks are articulated differently in various studies as "study skills" (Cheng, Myles, & Curtis, 2004), "study areas" (Woodrow, 2006); "tasks" (Banerjee, 2003); "coursework activities" (Cotton & Conrow, 1998); "course-related tasks" (Elder,

1993); “problems” (Ferguson & White, 1994); “difficulties” (Banerjee, 2003, pp.445-449; Cotton & Conrow, 1998, p.99; Ferguson & White, 1994; Weir et al, 2013); “language tasks” (Ingram & Bayliss, 2007); “academic reading and writing experiences” (Weir et al, 2013); “preparedness” (Byrne & Flood, 2005, p. 117) and have been articulated as “communication strategies” in the CEFR (Council of Europe, 2001, pp. 24-26). What they all articulate are situations or tasks in which students perform in the target language or areas in which they might encounter difficulties due to a lack of skill which can be exacerbated if language proficiency is also lacking. In one sense, the criterion of performance in academia is an abstraction:

> First, as Messick pointed out, the criterion is itself an abstraction, or construct. So, for example, in communicative language testing the target of test inferences is a set of performances in a particular context of use, which are in principle unobservable (McNamara, 2003, p. 467)

An example of this “unobservable” aspect of performance is the cognition of a student as they negotiate all the tasks and situations at university. There is also the issue in observing performance being misleading or inaccurate, for example a researcher can sit in a seminar room and observe the academic performance and linguistic behaviour of a test taker, but this could be problematic in how the student then performs while being “watched”. This situation can lead to phenomena such as performance anxiety or the “Hawthorne Effect” (Landsberger, 1958), whereby an individual worsens or improves their behaviour as a result of being observed.

Whilst it is true that direct observation of a student’s language behaviour may be somewhat artificial and inaccurate, and therefore “unobservable”, there are ways to assess performance in the criterion domain of university which have been carried out in previous studies. A student can, for example, self-assess their linguistic and academic performance in a range of tasks and situations at university. The methodology of these studies is investigated in the

review of studies below. Before doing this, I also mention the role of academic literacy in performance because this is anticipated to form part of the students' experience of academia in strand 2 of the research.

As stated above, performance in academia is on one level, about having the language skills with which to communicate and interact effectively with other students and tutors in a range of academic tasks and situations. At another level, it is about becoming familiar with the discipline you are studying in, the writing styles and genres, and acquiring the skills with which to study effectively at university – these may be referred to in the literature as "academic socialisation" (Lea & Street, 1998, 2006), or "all categories of EAP" (Wingate & Tribble, 2012, p.487). Again, these skills are needed by all students at university and are not generally taught or used in high school systems and are usually acquired at university.

Writing is a way to measure performance, or to gauge whether a student is coping in academia in their particular subject. On most academic courses, work can be assessed in various ways and forms, but much performance in academia is typically measured in terms of written or oral assignments as coursework; longer pieces of assessed writing such as dissertations and theses; practical assessments, and end-of-first term, end-of-year or final exam scores. Academic writing is therefore a concrete way to conceptualise and measure "performance" and very much involves being able to produce pieces of work which need a certain level of proficiency in English but also need to take account of the definition of proficiency encapsulated in "academic socialisation" and "academic literacies" – for example, following subject-appropriate register and style, using appropriate discipline-specific vocabulary and engaging with the discourses of the discipline and meeting tutor and department expectations of what is considered writing in a particular discipline.

Although it includes academic and linguistic skills, the term ‘academic literacy’ has a subject-specific application and is concerned with proficiency and skills which include but go beyond general academic language proficiency. Language proficiency in tests for academic performance refers to sufficient language for communicative competence needed in all subjects. However, even if a student performs well on an academic English test and can be said to be “ready for academia” in general linguistic terms, that is very different to being ready for the challenges of academia that await them in their particular subject.

2.2.2.4 Concluding Remarks

Achieving success in academia or ‘performing’ well is often a *combination* of subject-specific literacy skills (knowledge of subject-specific vocabulary and genres or styles), possession of good general study skills (the ability to cite effectively or construct a coherent argument) and a good level of language proficiency with which to do these. Even though they possess a certain high level of language proficiency, “native speakers” may not necessarily have the other skills mentioned here and so good academic performance and “success” in academia is not guaranteed simply by being proficient in the language.

This study is concerned with whether the scores from the PTEA can be interpreted for students from outer circle backgrounds as to whether they have sufficient language proficiency to be able to cope with the demands of university study *in English*. This is the main aim of the predictive validity study. However, I also anticipate that in interviewing students and gauging their impressions of performance, some issues may be to do with sufficient or insufficient language proficiency while others may be also to do with a deficit in other skills to do with “academic socialisation” and “academic literacy” (Lea & Street, 1998). Whether the tests indicate if students are “literate” for their particular subject is a question

that the academic tests cannot answer. In other words, a student may be "ready", "skilled" and "able to cope" in their L1, but the major question in my study is to do with whether they are ready for performing in the target language domain in university – in English – based on their proficiency test scores. This is theoretically indicated in the PTEA score report by the four *communicative* skills of reading, writing, listening and speaking, and by the six enabling skills scores.

Performing, or coping 'well' academically is not just dependent on having an appropriate level of language proficiency. On one hand, it has been measured in terms of "outcomes" by use of test scores, grades, grade point averages and others. In this regard it is dependent on academic literacy, because it often involves assessment of discipline-specific written work. On the other hand, performance is also defined and measured by the reflections, opinions and judgements of tutors, students and others as to whether a student is "coping". In this vein, performance in the academic domain is also about demonstrating skills and the ability to complete tasks in various contexts (lectures, seminars, tutorials) and also depends on variables or affective factors such as mental health, psychology, well-being, financial security and an aptitude for learning. These and other factors are discussed in the next section.

In a holistic conceptualisation of performance, a predictive validity study needs not only to look at final academic grades (which are supposed to summarise the achievements of a student in numerical form) but should also look at the actual skills and language behaviour of the student when they are on the course itself, using constructivist approaches such as gauging the perspectives of the students and other stakeholders of their academic journey. It is the job of academic English tests to gauge whether a student has sufficient language proficiency to perform adequately at university level in the language of instruction and to provide scores that can be interpreted in terms of a readiness for the language demands of

academia. However, predictive validity studies must each define what "performance" is in the target domain, and how it is to be measured. I now analyse the specific findings of past predictive validity studies to see how they conceptualised and measured performance, how they ascertained predictive validity and how they inform my study.

2.3 Findings from Predictive Validity Studies

2.3.1 Proficiency and Academic Performance in Past Predictive Validity Studies

There have been a number of predictive validity studies that have sought to ascertain a link between test scores and subsequent performance and academic success. However, in seeking to gather evidence about proficiency and performance in academia, predictive validity studies have largely been accompanied by the same recurring characteristics or caveats they did when Graham (1987, p. 506), wrote her paper nearly thirty years ago and the following three common aspects of studies have been identified in previous studies (Elder, 1993, Cotton & Conrow, 1998; Ferguson & White, 1994).

1. Defining proficiency

Firstly, there is disagreement about what is meant by proficiency and whether the measure of English proficiency is valid. In other words, there are contending views as to what language proficiency "looks like" not least as expressed in the form of competing test constructs, content and test scores. In my study, the PTEA *is* essentially the primary definition of proficiency I am working with since the study is a predictive validity study of this particular test but evidence for proficiency is also collected in the domain, the future criterion.

2. Measuring academic performance

As alluded to in the discussion above there is a lack of consensus about what exactly determines academic performance (the criterion). For example, it can be asked whether a performance, and by extension academic success, is only realised in final academic grades (outcomes) or whether it is measurable in other ways such as language behaviour and demonstration of skills and performance in tasks. Different predictive validity studies use different criteria.

3. Interpreting relationships

Thirdly, how should any findings of a relationship between proficiency and academic performance or success be interpreted, seeing that a lot of studies generally conclude that the relationship between proficiency in English and academic outcomes is "more ambiguous than one might initially suppose" (Cotton & Conrow, 1998, p. 75). Contributing to this difficulty in interpreting a relationship is the "time lapse problem" mentioned by Ferguson & White (1992, p.17) whereby there is a long gap between the taking of the proficiency test and the measurement of the criterion for academic performance in the domain of university. As much as one year or more may have passed between the two measurement points and therefore many factors may have come into play in the interim which affected the proficiency. This aspect is unavoidable unless for example every single test taker in a study takes the test a month or less before being accepted into university, or post-entry to the university in the first month of their studies. This is not feasible in many studies and as long as some tests have a two-year validity on them, this time lapse factor in predictive validity studies will always exist. Some studies (e.g. Humphreys et al 2012) have managed to control this factor in providing financial and organisational assistance to set synchronised academic tests to participants at the beginning and end of their studies but this was not an option in this study.

Connected to this is the existence of 'multiple variables' or factors that not only affect a student's *test performance* (such as preparation time, familiarity with format etc.) but crucially also the student's *academic performance* (such as financial hardship, psychological problems, workload and time pressures).

The following review of the methods and findings in previous studies reveals what other studies have done, conceptualised and discovered in the light of the three caveats. In contrast to the largely quantitative studies looking at a very particular measure of academic achievement, the studies below look more at "study experiences" or what I term "experiencing academia" at university as a way to conceptualise academic performance. The first section looks at how some studies have tried to assess language proficiency in academic performance.

2.3.1.1 Self-assessment (self-rating) of Language Proficiency

The idea or method of self-assessment or self-rating of linguistic proficiency and academic performance has been adopted and measured in a number of studies including Weir et al (2013); Ingram & Bayliss (2007); Fox (2004); Woodrow (2006); Hill et al (1999); Cotton & Conrow (1998); Ferguson & White, (1992) and Elder (1993). This method typically uses rating scales as a way to quantify opinions and is useful in predictive validity studies because it provides triangulation in respecting the participation and perspectives of the students themselves in being able to rate their own linguistic proficiency and performance in various aspects of academia.

In surveys, Hill et al (1999) asked international students at the University of Melbourne to rate their English proficiency using a scale from 1 = 'excellent' to 5= 'weak'. However, in this rudimentary survey using an odd number of choices with a middle value, students tended to rate themselves in the middle point of 3 (which may be a typical response in a scale using a middle choice) rendering that aspect of the survey of limited interpretability. The results were presented in terms of differences in ratings between those who had had sought ESL support (slightly higher proficiency) with those who did not, but the differences were not statistically significant.

Ferguson & White (1994) employed a similar self-rating instrument for international students at the University of Edinburgh using a scale of 1-7 between 'excellent' (7) and 'very bad' (1) for indicating proficiency in the overall and four skills in English. However their instrument could be considered more robust than that of Hill et al (in terms of comparability with other studies) as the scale descriptors were written to be comparable to the IELTS descriptors. Tutors also used the instrument and both measures were compared with the test takers' actual IELTS scores on two occasions after two separate IELTS tests. The Spearman statistic they obtained for the first test was 0.89 (indicating a "very strong" correlation) for the overall score representing "correlation of judgements to IELTS score" (p.32). However, this might not be considered valid as the rating was done 6-9 months after the IELTS test. The second correlation with a second IELTS test (close to each other in time) was at .57 (a "moderate" correlation). Correlations for student self-rating were low in comparison to that of tutors (.32 for the first IELTS test; .49 for the second). In a comment on the actual use of the self-assessment tool, Ferguson & White (1994, p. 34), suggested that weak students tended to overrate their language proficiency whilst stronger students tended to underrate themselves.

Another study by Ingram & Bayliss (2007), used a similar measure of publicly-available IELTS descriptors to promote self-rating of ability "The questionnaire asked students to rate their own language behaviour in reading, writing, speaking and listening by selecting one description from a range of six options for each skill area" (p.9). In their study of 28 students from undergraduate to PhD level at two Australian universities, they did not use statistical correlations and instead introduced the concept of "language behaviour" into their predictive validity study. Their instrument for self-assessment of proficiency revealed a wide variability in how students self-rated themselves with 36% of respondents rating themselves higher than their actual IELTS scores suggested, 25% rating themselves as equal, and 37% rating themselves lower. This could suggest that students are not very accurate at rating their own proficiency, for example, Kruger & Dunning (2009) talk about "inflated self-assessments" in that people are not always able to accurately rate themselves in their own abilities often due to being unskilled or unaware. However, as with Ferguson and White, Ingram & Bayliss's 2007 study reveals that students do not necessarily always over-rate themselves. Equally, there may be a valid disagreement or mismatch in how the proficiency test rates students' proficiency and how test takers view their own levels of proficiency. A consideration may be that while self-rating, a student may be assessing their general English ability whereas the test measures their ability from an academic construct of proficiency. One problem in Ingram & Bayliss's study was that the tool may have been compromised as the instructions to participants were to rate their language behaviour based on their interactions with "native speakers" but this is arguably highly dependent on the student's perception of who a native speaker is and indeed may be irrelevant in highly internationalised universities of today.

From these studies, self-assessment of proficiency provides useful triangulation with other measures but doubts, or rather a healthy amount of caution, should remain as to how

accurately students use any such self-assessment tool not least because of possible factors such as hurried responses, automatic "box ticking" and not reading descriptors properly.

2.3.1.2 Self-assessment of 'experiencing academia' in the Criterion Domain

Many previous predictive validity studies and studies of academic performance also gathered self-assessment data in the form of participants' perceptions of *performance* in the criterion domain in the tasks and situations discussed (section 2.2.2.3.2). For use in this study, I term this perception of performance 'experiencing academia' which aligns with the concepts of "real life" and "authenticity" (Bachman, 1990), in measuring students' perceptions of their actual language behaviour and/or their performance in a number of scenarios and tasks at university. Examples of where this concept has been used include studies such as Weir et al (2013); Ingram & Bayliss (2007); Woodrow (2006); Byrne & Flood, (2005); Cheng et al (2004); Banerjee, (2003); Hill et al (1999) and Ferguson & White, (1994). Typically, certain situations in the criterion domain of academia are presented to students such as "comfortable in asking for help from university lecturers" or "at ease working in groups" and the students are then asked to indicate their proficiency in these situations or to what extent they feel prepared for these situations often using quantitative measures such as Likert scales and calculation of proportional or mean responses.

Hill et al (1999) undertook a comparison of IELTS and TOEFL as indicators of academic success and looked at 35 IELTs test-takers and 27 TOEFL test-takers from 17 different L1 backgrounds. They used surveys and interviews to separate students into "academically successful" and "academically unsuccessful" students. Survey responses were entered into Excel and categorised (coding). Interviews were transcribed and comments categorised into

language related factors (proficiency), study-related factors (academic skills and academic literacy) and acclimatisation factors (non-linguistic). The survey included a tool to assess English as a Second Language (ESL) support. Students could comment on whether they had received any help and also rated academic skills in terms of difficulty. Using this initial survey method Hill et al identified student issues with "understanding and communicating (both formally and socially) with native speaker students; listening and note taking; reading (including reading speed); and writing (including problems with expression, style and organisation)" (p.70). In terms of non-linguistic factors impacting on students they discovered issues to do with "settling in" such as finding accommodation and adjusting to academic culture and tutor expectations. Students cited academic skills such as "difficulty in critical thinking"; "understanding concepts" and "dealing with workload". 22 students were identified for interviews in order to follow up survey responses in more detail. In these interviews students corroborated what had been mentioned in the survey as well as mentioning other factors which they perceived as affecting their grades. These included:

- Inadequate background knowledge (including cultural assumptions and knowledge of local conditions)
- Poor study skills or time management
- Difficulties in adjusting to a new culture and style of education
- Insufficient application (effort)

Language related problems cited included:

- Difficulty in following lectures
- Understanding native speakers in discussion

- Lack of familiarity with the relevant genre (i.e. for written assignments)
- Difficulty with reading speed
- Difficulty in completing assignments (seem to be at least partly linguistic in nature)

Weir et al (2013) looked at the GEPT (General English Proficiency Test) - a test introduced in Taiwan by the ministry of education in 2002. They compared GEPT reading and writing scores (independent variable) with self-assessment of academic reading and writing abilities and "real life academic performance" in writing tasks (course work and examinations) (ibid, p.1). They used Likert scales which allowed students to self assess their academic reading and writing abilities. They presented 29 different statements to participants that described different aspects of reading and writing and students could rate from 1-4 as to what extent they agreed or disagreed with the statements. This method revealed areas in which students were least confident and most confident. In reading, the least confident areas included "understanding long and complex sentences", "all vocabulary items" and "concern about time limits in reading tasks". Students were most confident (reading) in: "understanding instructions and visuals" and in their "careful reading abilities". In academic writing most problematic were: "considering and communicating to the intended reader"; "writing about complex subjects"; "using a range of vocabulary" and "editing for linguistic accuracy" (their own texts). They were most confident in: "finishing the writing task on time"; "presenting and supporting arguments well"; "macro-planning"; "referencing" and "organising ideas".

The authors added another quantitative element to this "real life" academic performance by comparing GEPT and IELTS reading and writing test scores with performance on "real life" academic tasks based on course materials at the researchers' location at the University of Bedfordshire: "two writing assignments (essay and report), one in-class test and one end-of-

term examination from four different modules" (p.24). They found that "reading scores correlated with the real-life academic performances at .499 ($p<.01$) and GEPT writing scores at .294 ($p<.01$)" (p.35) indicating a moderate positive and weak positive relationship respectively. However, the real life performances of students were only measured by one rater using an internal system of a scale from 0-16. The reliability of this can be called into question as there is no way of checking how reliably marked these real life performances were. This remains a problem in predictive validity studies where each institution has its own marking system for assessing work and where some work is marked by multiple assessors and some only by one. An additional problem is that the course materials and rater scale used as a comparison with the GETP scores were from the UK and not Taiwan.

An innovative methodological approach to looking at predictive validity was used by Banerjee (2003) at the University of Lancaster following her involvement in a preliminary study (Allwright & Banerjee, 1997). The preliminary study aimed to collect evidence of performance from 38 postgraduate students ('self-selected') with low levels of proficiency and the consequences for the university of admitting them (the "cost"). Although all were admitted to university, 29% of the sample had not met the university language proficiency requirements so Banerjee argued that the sample was 'less truncated' than in other studies. Allwright and Banerjee collected data including IELTS scores, two self-assessment questionnaires, tutor questionnaires and interviews of selected students' tutors, student interviews, collection of grades and degree marks. Questionnaires focused on how students thought they were coping with the linguistic demands of the course, their academic skills and particularly whether they felt their grammar and vocabulary were sufficient for academic study. Interviews focused on students' perspectives of "cost" - the discrepancy in proficiency between UK English native-speaking students and certain overseas students manifested in

terms of how much extra work and language support a student needed to undertake in order to be able to cope on their course (Banerjee, 2003, p. 253). In other words if a student needed to access linguistic support in any way differing from that of a "native speaker", this could be considered a "compensatory measure" which "costs" students time and effort.

From the interviews and questionnaires, despite the sample of international students consisting entirely of non-English L1s, non-linguistic factors were cited much more than linguistic factors. For example, the time pressure of having to produce a lot of work within deadlines was a common complaint. A lack of background knowledge or "topical knowledge" (Bachman & Palmer, 1996) and "unfamiliar teaching methods" were also cited. Tutors meanwhile cited a "lack of background knowledge", "personality" and "conceptual ability" as important factors affecting performance and rarely mentioned linguistic issues in isolation. The interviews, focusing on "cost" did however allow probing of the issues in the surveys and revealed that studying in an L2 such as English may well have contributed to students' feelings of time pressures (mentioned above) as it increased the time and effort needed to read, and to complete assignments as well as affecting their organisational ability in meeting deadlines and their linguistic performance in being unable to form complex ideas quickly in English. However, one problem with the study results lay in the wide score ranges of the participants ranging from 5.5 to 8.0 in IELTS. This fact coupled with the fact that *all* participants mentioned the problems above, made it difficult to ascertain whether the problems they experienced were attributable to having low language proficiency, studying in an L2 in general, or a lack of conceptual ability and subject knowledge (p.99).

Banerjee's later 2003 study used an ethnographic case study approach with methods mainly including interviews and student diaries based on detailed descriptions for 8 cases (3

Chinese-speaking, 3 Spanish-speaking and 2 other Asian L1-speaking participants), again at the University of Lancaster. Banerjee used three Lancaster University admissions categorisations for her cases, of "the clear accept", "a safe bet", and "a risk" (p.359), in order to assign the participants with one of these labels and then track their academic performance to see if she could provide evidence for those categorisations, rooted in actual student experience. For example, the student admitted to university under the heading of "risk" was theoretically most likely to experience "cost" due to language difficulties and a student categorised under the first heading of "the clear accept" was theorised as far less likely to experience problems due to linguistic proficiency but rather, other factors such as adjustment to the UK academic system for example, or study-skills related factors particular to that student. In general, Banerjee found that adjusting to British culture had a "tremendous effect" on the study experiences of her participants (p.388). On the surface, this is clear evidence for non-linguistic factors affecting performance. However, her two "clear accept" participants did not find this adjustment a particularly onerous task. Banerjee concluded that this was due to their flexibility linked to their language proficiency and past experience of other cultures. Banerjee identified the lower proficiency students' identification of cultural factors being perhaps used as a "scapegoat" for harder-to-identify *linguistic* issues (p.389). Overall, despite the categorisation of her participants under those three labels above, they all performed relatively similarly, academically-speaking. This led Banerjee to her conclusion that a direct linear link between language proficiency and academic grades is far from possible and she concluded that language proficiency is not a good predictor of academic success although care should be taken when generalising from case studies.

In addition to the self-rating of proficiency mentioned above, Ingram & Bayliss (2007) designed a semi-structured interview on student and staff perceptions of the extent to which

students were coping with their study demands, particularly in spoken and written performance. The 28 students in the study based in two Melbourne institutional campuses came mainly from Asian countries and all had IELTS score profiles. To provide further evidence of language behaviour the students were also observed in the academic domain (including being taped and videoed) in "class seminars, group discussions, oral presentations, lectures, and both in-class and out-of-class interactions with lecturers and other students" (ibid, p.11). They used language behavioural descriptors aligned to IELTS levels to rate student performance and behaviour in these areas and to compare with the participants' actual IELTS scores. As mentioned before, the reliability of observations of this kind could perhaps be queried (Landsberger, 1958; McNamara, 2003) because being directly observed and recorded by a researcher might have affected the performances. This aspect was later conceded by the authors as an uncontrollable variable intruding on performance as "levels of embarrassment encountered by students due to researcher presence" (ibid, p.14) and led to them not using statistical correlations of interview and observational data but instead using a discursive approach in describing the data. This produced a discursive description of features of language behaviour such as: "syntax, language functions and tasks, content and meaning, fluency and coherence, pronunciation, range of lexis, organisation of information, class involvement, pragmatic awareness and register" (p.14). The language behavioural descriptors used to rate each student in each academic context were aligned to IELTS levels. This method allowed the researchers to compare "behaviour in the academic context to the language behaviour implied by the students' actual IELTS test scores" (p.17). They found that 25 out of the 28 students (looking at overall scores) were rated at a level "which suggested language behaviour that equalled or exceeded that implied by their overall IELTS rating" (ibid).

From the student interviews, Ingram & Bayliss were able to gauge the confidence levels of students in various university scenarios such as "listening & note-taking"; "following spoken instructions"; "following written instructions" and "group discussions/tutorials". Students' levels of confidence expressed in numerical form were tabulated against their IELTS scores and what they found was that there was no relationship between IELTS speaking scores and "confidence in speaking" in the scenarios they had operationalised in the questionnaires "only at a speaking proficiency level of 8.0 did students feel completely confident in all speaking situations" (2007, p.23). They later discovered that students' concerns about being accurate had led to a decrease in confidence when speaking with tutors and lecturers. Students were most confident in English in non-academic scenarios around campus such as the cafeteria where accuracy was not such a concern. In terms of listening, no relationship was found between students' listening scores and their self-reported experiences on campus. Difficulties in academic settings were usually attributed to vocabulary. However a significant percentage of respondents (14%) cited social conversations as difficult, especially due to encountering different accents and idiomatic use of Englishes. Writing scores in IELTS did not correlate to students' responses in interviews but "14% said they found writing tasks extremely difficult and a further 43% found them quite difficult". Thus, writing assignments seemed quite challenging for many of these students. Specific problems with writing included a "lack of vocabulary"; "the ability to paraphrase; "use of grammar; "academic style" and "knowledge of the topic". Reading skills had a similar lack of relationship to IELTS scores in the student surveys although students with very low IELTS scores (5.5) did admit to rarely understanding reading materials. As for lecturers, none of them identified any lack of confidence in students and they perceived students to be "more participatory over time" (2007, p.36) though some students were difficult to comprehend in terms of spoken English (limited vocabulary, pronunciation and grammatical problems). Writing problems included: "lack of content and

limited ideas, lack of depth, grammatical inaccuracy, inability to express ideas in students' own words and organisation of information" (2007, p.38).

Fox (2004) looked at "test decisions over time" as an alternative model of predictive validity away from correlation of test scores with GPA. She looked at cohorts of largely expanding circle students, Mandarin, Cantonese and Arabic speakers from East Asia, the Middle East and North Africa in a Canadian study. All participants had taken the CAEL (Canadian Academic English Language) Assessment and the study included interviews with them and their EAP tutors, questionnaires, test performances and performances in ESL/EAP and discipline-specific classrooms. All were engineering, computer science or business UG students. The study sought to explore whether any of the students had been "misplaced" as a result of the test and therefore were "at risk" or struggling, a similar theme to that of Banerjee's 'cost'. Fox's study identified that students at risk often had a superficially fluent English (p.454) but on closer inspection struggled with the academic content and tasks in the courses. This was also due to a lack of "study strategy problems" such as working on their own initiative, and due to an "inflated sense of their abilities" (ibid, p.451) a concept also reflected in Kruger and Dunning's paper (2009). Out of 341 students in the programme the CAEL was considered an effective test in identifying EAP support. 15 of the 341 students (15%) were considered "at risk". In terms of factors affecting academic success the study identified "acculturation factors" as being important, for example adjustment to the society and culture, (including that of the university) was important in determining academic success. Fox's study indicates that it is valuable to investigate predictive validity as "the utility and impact of test decisions over time" (Fox, 2004, p.461) as an alternative model to the traditional correlation predictive validity studies.

Woodrow (2006) looked at "study problems experienced by students" as perceived by students and tutors. She used a questionnaire with a seven-point Likert scale and asked students and staff to rate *perceived difficulty* of course units (from 1 = easy to 7 = difficult) and *study problems* (from 1 = no problems to 7 = problems). These were quantified and correlated against GPA but no significant correlations were found. She looked at a limited range of issues of language, content and resources and found that:

> "The students reported finding resources and assignments to be the most challenging, with the content and language of assignments being equally difficult. The staff sample reported assignments being the biggest problem facing students with the language aspect being the most problematic." (p.65).

However, she did not investigate non-linguistic variables affecting academic performance which may have provided an extra dimension to her study. She did, however, have open-ended questions gathering information on: "perception of the adequacy of English proficiency measures to predict academic success"; "previous professional and academic experience" and "previous English language learning". Woodrow's conclusions were that academic achievement is only partly based on language and is a "complex issue". She recommended more communication between staff and between staff and students in terms of discussing the expectations of academic study.

Ferguson & White, (1992, 1994) designed a survey for a sample of 28 Masters students of various nationalities on life science courses at Edinburgh University. IELTS scores were collected, then participants were re-tested on IELTS in the following June. Both scores were then compared with their academic results. Student and supervisor questionnaires were also completed which had a number of interesting facets, including student perceptions on success, perceptions of problems on the course, advice to future students and self-reports on language difficulties. With regards the findings, most of the problems were initially not thought of as language-related as such, but were course-related. For example, "workload"

ranked as the number one problem together with angst about the "intensity and speed" of a Masters programme and questions around being "organised" – by both home and overseas students. However, it was suggested that language problems may have been due to the intensity of the course and again the speed needed in reading and writing at this level. Personal problems affecting performance were uncommon.

Other studies focused on academic performance in terms of adjusting to a different academic culture as a wider theme. Brown (2008) undertook a longitudinal ethnographic study of 13 international students over a year in their studies in the UK. She interviewed them 4 times over the year and conducted informal conversations. She identified essay writing as one area where students were unfamiliar in terms of writing skills such as "structuring essays", "paragraph formation", "referencing conventions", and "use of academic vocabulary" (Ibid, p.30). The need to develop critical thinking skills was also apparent. In the same study, managing a workload was also cited as a source of stress for international students and using "learning resources" in particular having to adapt to using computers as the main way to conduct academic work was stressful for the Russian and African students in the study (South African, Ghanaian and Nigerian). Similarly, Cheng et al (2004) conducted surveys (n=59) and follow-up interviews (n=12) with international students in a Canadian institution and in common with Brown, identified writing skills as posing the most problems to the participants in their study. Students with English as a non-native language found dealing simultaneously with language and academic demands to be a frustrating experience and so there was a recommendation for targeted support for oral presentations and written assignments.

Adjustment to academia was also a concern for L1 English speakers of the inner circle. Bryne & Flood (2005) conducted a study following concerns about student attrition rates in Ireland

and investigated accounting students' preparedness for HE. They discovered that although confident in their ability for most "learning tasks" at university "fewer students are certain about knowing what is expected of them academically or in their ability to work in groups or independently" (Byrne & Flood, 2005 p.118) and fewer still were confident about use of computers.

In terms of how these studies all inform my study it is apparent that it has been difficult in previous studies to identify separate linguistic problems from academic skills and academic literacy issues. Having sufficient language proficiency obviously plays an important role in adjusting to and acquiring other skills needed in academia but it is often difficult to isolate student problems in adjusting to academia solely in terms of a lack of proficiency as a cause. To add to this mix of variables affecting performance, an aspect which also plays its' part is that of the unknown variables. I now give a brief overview of the variety of these factors that have been picked up on in studies.

2.3.1.3 Multiple, 'Confounded', Intervening Variables

Quantitative, qualitative and mixed methods studies have suggested that there are a variety of intervening variables or "confounded" uncontrolled variables that affect academic performance. "Confounded variables" (Kish, 1959) are ones which interfere with the "perfect" model of the theory that there is a direct link between language proficiency and academic success or performance. Confounded or confounding variables in studies such as this would include those which are not controlled for in participants and include a "cocktail of factors" (Ferguson & White, 1992) such as financial difficulties, mental and psychological factors and academic ability.

Banerjee's study revealed that admissions were often looking for a combination of such variables that might indicate a "potential to succeed in the course" (2003, p.236). These were "language proficiency, academic background, intellectual capacity, work experience and personal characteristics" (ibid, p.236-237). It is difficult however, to prove to what degree they affect performance. At the very least, the studies above have all suggested a number of "intervening variables" which *contribute* to academic success beyond linguistic proficiency and *impact* on a student's academic performance. I include the variables below under categories in order to illustrate the wide variety of factors that have been identified in previous studies as potentially impacting on academic performance.

a) Possession of academic skills

These factors include: "lack of academic ability dominant" (Ferguson & White, 1992:61); "intelligence, academic knowledge and ability" (Davies, 2007 p.82); "study habits" (Ferguson & White, 1992); "time management and consistency" (Hill et al, 1999); "scholastic aptitude" (Elder, 1993); a "lack of reading skills and background knowledge" (Kerstjens & Nery, 2000), having to demonstrate "critical thinking" (Brown, 2008) and "self-regulatory learning strategies" (Karimi, 2010). Connected to the above is the wider issue of adaptation to the academic and cultural context (see b, below)

b) Adaptation or Acculturation

These factors range from: "social and cultural adjustments, learning & educational styles" (Kerstjens & Nery, 2000); "cultural adjustment" (Cotton & Conrow, 1998); "cultural differences" and "adjusting to educational environment" (Hill et al, 1999); "poor acculturation" (Fox, 2004); "understanding of classroom role relationships" (Elder, 1993), "anxiety about classroom discussions" (Brown, 2008); "independence and freedom at

university as a challenge" (Byrne & Flood 2005); use of "humour" in the classroom (Nesi, 2012) and the general challenges of transition to university whether in a domestic tertiary setting (Byrne & Flood, 2005; Karimi, 2010) or in one abroad (Andrade, 2006; Hyams-Ssekasi, 2012) including a lack of knowledge of what is expected of university students (Woodrow, 2006; Byrne & Flood, 2005; Hill et al, 1999).

c) Background factors (nationality, L1s, age, gender) including previous experience in English range from: "amount of English language tuition received" (Cotton & Conrow, 1998) to "ESL vs. EFL background" (Zheng & Wei, 2014); "prior educational achievement" (A. Duff, Boyle, Dunleavy, & Ferguson, 2004) to "past experience" (Woodrow, 2006, p.63).

d) Affective Factors

These range from: "affective and personal problems" (Ferguson & White, 1992); "extraversion, openness to experience, neuroticism, agreeableness" (Duff et al, 2004); "attitude" (Davies, 2007); "emotional problems" (Hill et al, 1999); "motivation" (Kerstjens & Nery, 2000; Cotton & Conrow, 1998); "maturity" (Kerstjens & Nery, 2000); "motivation and grit" (Bridgeman et al, 2015); "confidence" (Bayliss & Ingram, 2006; Byrne & Flood, 2005; Zheng & Wei, 2014); "anxiety, stress, homesickness and loneliness" (Andrade, 2006; Brown & Holloway, 2008) and "personality traits" (Karimi, 2010).

e) Other Personal & Financial Circumstances

These include: "family responsibilities" (Ferguson & White, 1992); "health" (Davies, 2007; Hill et al, 1999); "financial and family pressures" (Kerstjens & Nery, 2000); "welfare difficulties" (Cotton & Conrow, 1998); "finding accommodation" (Hill et al, 1999); a "lack of tutor feedback" (Cotton & Conrow, 1998) and "economic and practical concerns with UK living" (Gichura, 2010; Hyams-Ssekasi, 2012).

It can be concluded that there are a whole range of variables to be aware of that can impact on academic performance when conducting studies of this nature. The third category (c) is, however, of particular interest in this study as it focuses on the outer circle status and L1 of participants (background or test taker characteristics) so what now follows is a short review of those studies that have touched on this aspect of the background of participants before articulation of the research gap and research questions.

2.3.1.4 Nationality as a Variable in Academic Performance

With regards to category c) above, none of the predictive validity studies have used a conceptual framework of the circles of World Englishes (Kachru, 1985) but some have looked at test taker demographics in terms of nationality and L1 and a number of interesting findings occur.

In Ingram & Bayliss (2007) five of the 28 participants came from outer circle countries (Hong Kong, Botswana and Malaysia), of which four of those five test takers scored 7.5 and above in the IELTS test as compared to the other 23 expanding circle test takers in their cohort who scored between 5 and 6.5 (with the exception of two Chinese students who scored 7 and 7.5 respectively). This is a small indication of potential proficiency differences between the outer and expanding circles but no particular distinction was made according to nationality in the indices of academic performance in their study.

In Johnson's study (1988), Malaysian students (outer circle) had a high mean TOEFL score whilst Japanese students (expanding circle) had a low mean TOEFL. However, their GPAs were similar. This difference in test scores was mirrored in Light el al (1987) where Indian

sub-continent students, (India, Bangladesh, Sri Lanka and Pakistan) had the highest mean scores for the TOEFL as compared to expanding circle students in the sample such as Korean, Thai, Japanese and Chinese but their GPA was only slightly above the mean GPA for the study. It was also hard to ascertain other outer circle performances in their study because Anglophone Africa was lumped together with Lusophone and Francophone Africa while Malays and Filipinos (outer circle) were put together with Indonesians (expanding circle). Consequently, I cannot say that Kachru's circles formed a conceptual framework for this study in the presentation of the results or analysis.

Dooey and Oliver's study (2002) took 89 individuals of business, science and engineering backgrounds from 15 nationalities (mostly south-east Asian countries) speaking 13 different L1s. Most of the sample came from Malaysia and Indonesia, and the authors do make some interesting observations according to nationality such as the Malays tending to state English as their L1s but the focus of their study was not to compare the performance of English-speaking Malaysian students with speakers of other languages in Southeast Asia. There were some other outer circle students in their study from Hong Kong, Maldives, Brunei, Pakistan as well as some from those nationalities who had permanent residence status in Australia. The latter were treated in the study as native speakers and consequently excluded from much of the study. Interestingly, 15 out of 23 of these "home" students did not achieve a pass mark in their study reflecting the fact that language proficiency alone is insufficient for predicting academic success. Dooey and Oliver also state that the IELTS test was not designed for native speaking students "and it is clear from the current study that IELTS is not predictive for such prospective students" (ibid, p.51).

Ferguson and White (1994) took 28 participants from a variety of countries including 8 from the outer circle (1 from Malaysia, 3 from Bangladesh, 3 from Nepal and 1 from Sri Lanka) but did not make any comparisons based on nationality. They did, however, talk about the "composition of student group" in terms of non-native and native speakers of English. They found that the non-native and native composition of their sample was not significant in terms of who needed to adjust more to the academic culture at Edinburgh. They found that a European/non-European divide was more significant in terms of non-Europeans needing more adjustment to the academic culture at Edinburgh University regardless of whether they were from the outer or expanding circles.

Nationality is, however, a significant factor in one study by Zheng & Wei (2014) where they talk of "anxiety" among Chinese students in terms of speaking compared to a "linguistic confidence" among Indian students. Zheng & Wei suggest that linguistic confidence is an important variable in the performance of students at university and I suggest that this may be linked to their ownership of English. These researchers compared the PTEA scores of Indian and Chinese test takers and it was found that Indian students displayed higher score profiles than the Chinese students which the authors explain by the fact that "English is a foreign language in China, but an institutionalized additional language in India" (Zheng & Wei, 2014 p.139).

It is clear from this study that the researchers are making a distinction along the lines of nationality - the Indian students' relationship to English was deemed to be more along the lines of English as a Second Language (ESL) than the Chinese students' English as a Foreign Language (EFL) relationship to English. Even though the Indian students' lowest scores were for grammar, Indian students generally outperformed the Chinese students in all score categories on the PTEA and a t-test indicated that this difference was statistically significant

in all score categories except written discourse (ibid, p.129). Furthermore, classroom anxiety and speaking anxiety were two significant predictors for the Chinese but not among the Indian students (p.139-140). Zheng and Wei specifically put this proficiency difference down to differing affective factors such as the fact that one third of the Indian students in the survey said that people in their school spoke English. The authors also cited the differing teaching techniques of English in their respective countries contributing to either learner anxiety towards the test (China) or linguistic confidence (India), but don't mention the fact that English has a different history in Indian schools and society than it does in China (India being an outer circle country where English is more embedded in the society and education system). Zheng and Wei conclude that the EFL/ESL divide indicated by the differing score profiles in the sub skills has implications for providing different support and preparation for the test.

This difference in mean score test performance between outer circle Indian and expanding circle Chinese test takers was also found in a TOEFL study by Bridgeman et al (2015, p. 10) indicating that Indian students (TOEFL 103) had a higher proficiency in English than Chinese test takers (TOEFL 73). However, in this study the sample sizes were very small and Chinese students exhibited a higher GPA suggesting again that "academic success" is down to more than language proficiency.

In Gichura's study on Kenyan students and academic performance in the UK (Gichura, 2010) she stated that language proficiency was not a factor due to Kenyans competency in English (ibid, p.77). Instead practical problems such as economic concerns were identified as most pressing for them (ibid, p.260). Furthermore, acculturation issues affected them such as the food, weather, finding work and financial hardship "some of them experienced financial, social, practical and personal problems which resulted in negative study experience" (p.272).

Similar non-linguistic proficiency problems affected the transition experiences of African students in a study by Hyams-Ssekasi (2012) but in that study it was found that students did face issues with others' understanding of their accents and varieties of English including a degree of intolerance.

The findings from these studies indicate that there is a difference in proficiency between test-takers from the outer and expanding circles (as reported by mean test scores) which could reflect the differing backgrounds in English of these two groups. This provides a gap for this study.

2.3.1.5 Reflecting on the Previous Studies

The studies generally indicate that there are multiple ways of conceptualising and measuring academic performance at university. Some only look at "outcomes" focused on a final academic grade or average grades over a semester using statistical tests. Other studies looking at "experiencing academia" in terms of study difficulties and language behaviour show that qualitative and quantitative methods can be used to conceptualise and measure academic performance using a variety of instruments such as surveys, interviews, observations and self-rating tools. Although language is a significant factor in academic success, it is difficult to prove a direct linear relationship between linguistic proficiency levels in the form of test scores and academic performance expressed in grades and marks. It is clear from looking at previous predictive validity studies that those studies that only rely on a final academic grade as a measure of academic performance as criteria for comparison with the language proficiency score offer insufficient evidence for the validity or 'predictive power' of a test.

There are also multiple variables beyond language that affect performance, especially after a certain linguistic threshold has been reached. These variables such as adjusting to a new academic culture can affect both "non-native" and "native" English-speaking students alike.

In terms of study design, what these studies reveal is that there a problem in having separate survey self-assessment instruments for *linguistic* skills and *academic* skills. This is because in practice, it is often difficult to distinguish the two skills: "language and academic ability probably overlap; one tends to go with the other" (Ferguson & White, 1992 p.52). However, by measuring academic (including linguistic) performance in various scenarios as illustrated in the methodology of various studies above, it is useful to see exactly *which* aspects of language and skills are required to be able to function in a university. The methodology and findings of these studies are valuable in conceptualising "performance" or "experiencing academia" in the target domain of university and reveal that many factors affect the academic performance of students at university.

There is therefore a gap worth exploring not only in terms of the language proficiency of outer circle students as reported by language proficiency test scores but also in terms of a need to widen the definition of academic performance beyond only course grade or score-based "achievement". What follows is an articulation of the gap, which informs the current study.

2.4 The Research Gap

Firstly, in evaluating the literature, there has been very little investigation of assessing the English language proficiency of outer circle students as a conceptual group. In particular,

there is a need to explore how test-takers from this grouping perform on academic English proficiency tests, how they are admitted to university in the inner circle, and how their theorised *de facto* "native" or "native user" (Davies, 2013) proficiency relates to their actual experiences with the language demands and aspects of academic performance on their courses. A particular characteristic of the test takers in my study is an already acquired proficiency in English. This proficiency is explored using large-scale test score data and statistical techniques as well as qualitative interviews and surveys. Only a few studies have looked at outer circle students' actual test performance on academic English tests. Those that have looked at test performance (Zheng & Wei, 2014), have concluded that the nationality background of the students is a factor in their test performance, particularly regarding confidence in spoken English which I relate to having a certain ownership of the language.

Secondly, following studies such as Banerjee (2003), there is a need for more research into predictive validity outside of defining academic performance *solely* in terms of a calculation of correlation and regression analyses between final academic grades and language proficiency test scores. There has been very little research done with qualitative methods on predictive validity, especially regarding conducting case studies of students, and their perspectives on their educational and linguistic "journey" through the first year of academia. My research allows comparison of how their actual score profiles match with their actual linguistic behaviour and academic performance at university, what variables affect their performance, and what support is needed by these students in their academic life.

New directions in predictive validity studies are needed because: "the research model upon which predictive validity studies are based is unable to reflect the reality of the complex, multifactorial, dynamic process of learning" (Lloyd-Jones, Neame, & Medaney, 2012, p. 8)

In order to address this gap in the research, I conceptualise academic performance as more of a "process" of learning and adapting to academia and less of an "outcome" that is only quantifiable in a grade. There are a multitude of academic, linguistic, social and personal factors (in any combination) that influence the performance of a student on any degree programme. I aim to be able to capture this "complex" and "multifactorial" process in my research by using mixed methods – outlined in the following methodology chapter.

I define academic performance as including academic "achievement" in terms of grades, marks and final awards but also encompassing other aspects. I will employ research tools to allow me to gauge the student's perspectives on their language proficiency (self-rating); their ability to cope with the language for the demands of the course; the study problems that they face; perceptions of their performance in domain-specific tasks and situations such as writing essays and contributing in seminars. These aspects of defining and measuring performance have been described in previous studies: (Weir et al, 2013, Ingram & Bayliss, 2007, Woodrow, 2006, Cheng et al, 2004, Banerjee, 2003, Ferguson & White, 1994), but in my study I interpret the findings from the second strand in light of the large-scale quantitative data and qualitative data in the first strand which sheds light on the proficiency and ownership of English of my participants.

This study follows recommendations made in other studies to investigate other factors affecting academic performance. I combine this with the gap in the research of focusing on students who have hitherto been ignored as a conceptual grouping in predictive validity studies – those from the outer circle. I place my study in the tradition of tracking studies, case studies and studies on perceptions of the validity and usefulness of test scores already conducted by researchers cited above and others such as Lloyd-Jones et al, (2012), Fox,

(2004), Hyatt & Brooks, (2009), Humphreys et al, (2012), Coleman et al, (2003), Kerstjens & Nery (2000) and Cotton & Conrow (1998). This study adds to those predictive validity studies previously carried out, by focusing on an overlooked group of test takers – those from outer- circle countries with a proficiency in and ownership of English. The overarching aim of the study is to better support these students at university by looking at their overall experiences, and to explore how the test scores are interpretable for them. This now leads to my research questions.

2.5 The Research Questions

The research questions for the two strands of research are as follows:

STRAND 1 - PROFICIENCY and OWNERSHIP

Strand 1 looks at the English language proficiency and ownership of English of outer circle students for admission to university in inner circle countries.

Strand 1 Research Questions

RQ1 (ADMISSIONS):

What do the admissions policies of universities regarding nationality and CSELTs qualifications (as proof of English proficiency) suggest about the ownership of English and the English language proficiency of outer circle students?

RQ2 (PROFICIENCY DIFFERENCE):

i) Is there a difference in the proficiency of outer circle test takers compared to expanding circle test takers as reported by PTEA and other academic test scores (IELTS and TOEFL-iBT)?

ii) Is there a difference in proficiency, as reported by PTEA test scores, according to the L1 of the test taker? (English L1 vs. Other L1)

RQ3 (PERCEPTIONS OF PROFICIENCY-OWNERSHIP):

What are stakeholder perceptions of outer circle proficiency and ownership of English for university entrance?

STRAND 2 - ACADEMIC PERFORMANCE (Case Study)

Strand 2 looks at the relationship between the PTEA score profile and the subsequent academic performance of test takers from an 'outer circle' English background.

Strand 2 Research Questions

RQ4 (PREDICTIVE VALIDITY):

How do the PTEA score profiles of individual students compare to their actual language behaviour and academic performance in their first year?

RQ5 (VARIABLES):

What are the multiple variables that affect the academic performance of these particular outer circle test takers on their chosen courses?

Chapter 3: Methodology

3.1 Mixed Methods Design

3.1.1 Overview

In this study I adopted a mixed methods approach (Teddlie & Tashakkori, 2006). The particular design is of a convergent parallel mixed methods nature (Creswell, 2014, p. 77; Jick, 1979). A convergent parallel mixed methods approach to research allows the researcher to tackle the research in two "strands" which in the typology to date usually consists of a quantitative and a qualitative strand (Creswell & Plano Cark, 2011, p.63). The two strands or "halves" of my study involve collection of a number of data sets that combine to form the whole research story. The first strand in my study consists of quantitative and qualitative data and the second strand comprises largely qualitative data. In this study the strands are converged mainly at the theoretical framework and interpretation of results stage. Both strands of enquiry produce data that are analysed as per instrument and answer specific research questions within the over-arching theme of the English proficiency, and ownership of English of outer circle students for academia in the inner circle. Analysis of themes and interpretation of evidence using findings from across the data sets enhance the discussion.

The aim of the first strand of enquiry was to assess the English proficiency and ownership of outer circle students as reported by quantitative and qualitative data in the context of admission to academia. This was done mainly through analysis of a large data set of academic English proficiency test scores in the case of the PTEA data consisting of the scores of individual test takers and also publicly-available online data on mean test scores for IELTS and TOEFL-iBT tests according to nationality. The large quantitative data set from Pearson on their PTEA test data from 2014 (n=8067) allowed me to calculate mean test scores for the overall and four communicative skills and to categorise these by country and L1s within

countries and groups. This allowed me to gauge what variations exist in proficiency within and between the outer and expanding circle groups as reported by test scores. Additionally, a large-scale survey of publicly-accessible university admissions policies (documentary data) from 132 HEI websites allowed me to assess to what extent UK universities regard alternative outer circle qualifications (in the form of country-specific English language tests), as legitimate, pre-existing evidence of language proficiency. These two data sets are then complemented by two mostly qualitative data sets. Firstly, a survey of 17 outer circle students (online questionnaires) on their perceptions and views towards their proficiency in English together with their views on, and experiences of UK admissions policies regarding their CSELTs. Secondly, interviews with three Pearson country representatives on perceptions of language proficiency and ownership of English within three outer circle "markets". This combination of data sets offers fresh insights into a hitherto under-researched area on the test-taking experiences and attitudes of outer circle test-takers.

The overall aim of the second strand of enquiry was to take a case study approach to examine the predictive validity of the PTEA in terms of comparing students' language proficiency (as reported by test scores) with their perceptions of their proficiency and academic experiences. After conducting a recruitment survey of 46 students, four outer circle students from this survey became the "cases" studying within the UK (inner circle) and I conducted up to three tutorials with each of them over a year, collecting their PTEA score profiles and documenting their experiences in academia including analysis of their written assignments and their perceptions of their performance in the academic domain. I refer to them from now on as "cases". Their academic experiences were then compared with their score profiles to see if their PTEA scores had indicated any areas of linguistic difficulty that they had faced on their respective courses. The strands and the research questions linked to data sets are summarised and illustrated below and in figures 2 and 3.

3.1.1.1 Strands and Data

1st Strand

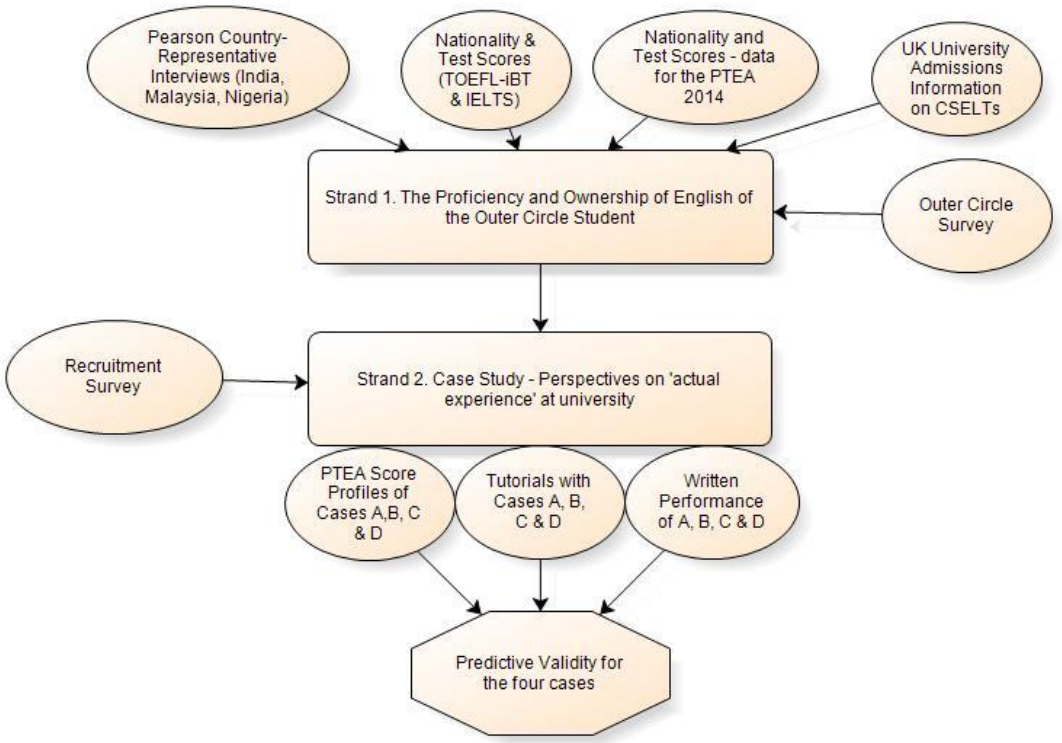
- University admission policies regarding CSELTs (132 institutions)
- Large data sets of PTEA test scores, nationality and L1s (n=8067)
- Publicly-available mean test score and nationality data sets from IELTS and TOEFL-iBT (number of participants not published)
- Stakeholder perceptions of outer circle proficiency and ownership: an outer circle questionnaire (n=17); and interviews with Pearson country representatives (n=3).

2nd Strand

- Recruitment Survey (n=46)
- Then a Case Study (n=4) and collating the cases': PTEA score profiles; self-rating of language behaviour; self-reflections on their "academic experience" on their courses; evidence of written performance (examples of assignments, feedback and grades).

The relationship between the two strands of the study is illustrated in the 'map' of the data in figure 2 below.

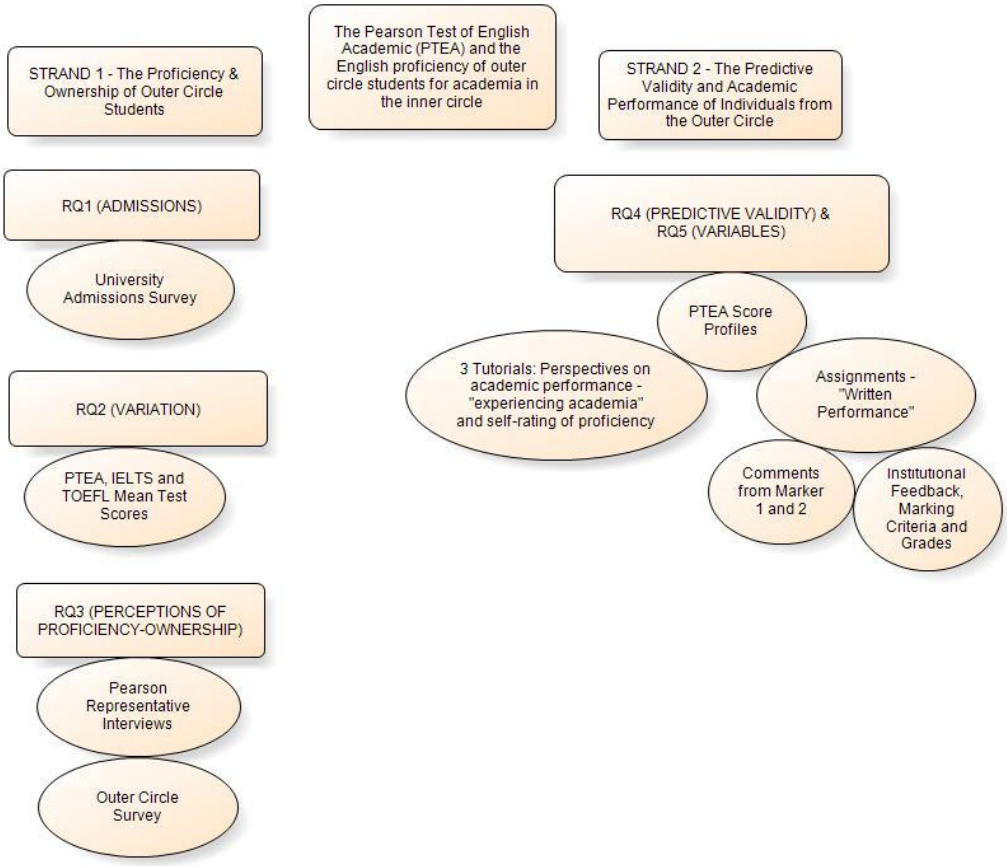
Figure 2: The 'Map' of the Data



The two strands in rectangles consist of the surrounding oval data sets. Strand 1 sets the background for the themes of proficiency and ownership of English, this then connects with strand 2 as I explore the actual academic experience and language proficiency of a group of outer circle students at university. The recruitment survey "recruited" four individuals to the case study. Analysis of their academic experiences led to discussion of the predictive validity of the PTEA for the four cases (octagon). The overall theme of both sets is the proficiency and ownership of English of outer circle students in academia and their academic experiences at university (their performance).

In terms of how the research questions fit into this design, the diagram below outlines which research questions are connected to which data set. The data sets are in the oval shapes underneath their respective research questions in rectangles. The overall theme of the research is in the centre:

Figure 3: The Research Questions and Data



The following sections describe how these data sets, research questions and strands fit together in the study design.

3.1.1.2 Level of Interaction of the Strands

The design of the study was fixed in that "the use of quantitative and qualitative methods is predetermined and planned at the start of the research process, and the procedures are implemented as planned" (Creswell & Plano Clark, 2011, p.54). The strands are independent of each other in terms of research questions and data collection, yet their theoretical design, analysis of certain themes and interpretation of results complement each other (Greene,

2007). The strands have 'levels of interaction' (Creswell & Plano Clark, 2011, p.64-65) in these three places:

1. The overall theoretical framework in the study is an exploration of the ownership and proficiency in English of outer circle students (section 2.1.5, p.57-60) for academia in the inner circle. Exploring this framework is the overall "programme objective" (Greene, 2007). This theoretical framework of the study is important for both strands of the research.

2. At the level of the data analysis, individual data sets and quantitative and qualitative data were analysed separately and the results presented separately. However, data such as the Pearson interviews and outer circle survey in the first strand produced themes and these were compared in the discussion to the themes emerging from the tutorials in the second strand largely in relation to the topics of proficiency and ownership of English in a "joint display of data" (Creswell, 2014, p.223). This comparison takes place in the discussion section. The large-scale quantitative data from strand 1 (test scores, Pearson interviews, and outer circle survey) is used to give a context to the data in the case study.

3. In the discussion of the data findings I interpreted the results and drew conclusions using all the data to answer the research questions and address the overall theme of the proficiency and ownership of English of outer circle students in academia in the inner circle:

> In the final step, the researcher interprets to what extent and in what ways the two sets of results converge, diverge from each other, relate to each other, and/or combine to create a better understanding in response to the study's overall purpose (Creswell & Plano Clark, 2011, p.78).

3.1.1.3 Priority & Timing

The two strands have an equal priority in the study; they have an equally important role in addressing the research questions. The qualitative research is not subordinate to the quantitative research; it complements it. The quantitative data gives context and background to the qualitative research - the case studies in particular focus on actual individuals as illustrations of outer circle students in academia. The timing of the data collection was not important in terms of the large quantitative data set (PTEA Scores) and qualitative data set (Case Study) as to which was collected first. They were collected at different stages in no particular order in relation to each strand. In this way my design was largely a "parallel mixed design" leaving out the word "concurrent" to emphasise that the qualitative and quantitative data collection did not necessarily occur at exactly the same time (Teddlie & Tashakkori, 2009, p. 144).

The Pearson interviews and outer circle survey were designed after the first tutorials of the case studies and so some themes and issues arising from those tutorials were incorporated into some of the questions within those research instruments. However, *within* the second strand, timing was important in that the PTEA score data for each participant was collected first and the three tutorials were conducted at specific times over the first year of study (see schedule below).

Table 1: Data Collection Schedule

Date	Comments
June 2013	A first 'sweep' of potential participants (PTEA test-takers) was conducted in a first recruitment survey mail out (n=41)
October 2013	A second recruitment survey was sent out but responses were few (n=5)
November 2013 – February 2014	Within this 'window' all the first tutorials were completed (n=4). Cases were labelled #A, #B, #C and #D
May 2014	By this date all the second tutorials were completed except case #A
September – October 2014	Interviews with Pearson Sales Representatives (Malaysia, Nigeria, India) were completed via Skype. A large data set was received from Pearson on country-specific PTEA scores.
November 2014 onwards	An online questionnaire of outer circle students was sent out to various university student societies nationwide (over 100 student societies).
December 2014	By this date the third tutorials were completed for #B and #C and examples of 3-4- assignments from each student (#B and #C) were provided as examples of written work.
December 2014 & August 2015 onwards	I received and processed a large quantitative data set from Pearson on test takers and score profiles from 2014. 8067 individual test scores were provided. University admissions information was sought from university websites

3.1.1.4 Bringing the Data Together

The challenge in a parallel mixed methods design is to bring the data together –a process of "convergence" or "merging" (Creswell, 2014, p.222). In this study I took a "side by side" approach (ibid, p.223) whereby each of the data sets above was analysed separately in terms of what they revealed about proficiency and ownership in the outer circle, students' test taking experiences, linguistic proficiency issues and academic experiences. In this approach, findings were analysed separately but then "joined" and compared in the discussion. Each data set answers a specific research question but after separate analysis I amalgamated findings from my instruments using an organisational method of "matrices" from Miles and Huberman (1994; 2014) in order to "bring the data together" in order to compare findings in tables and assist the final discussion. This idea of matrices is a way to display the data in a

visual (in this case a table) from a variety of sources and is basically an "intersection of two lists, set up as rows and columns" (1994, p.109) to display data from a variety of sources. Miles et al. argue that this technique avoids the "selective stacking" in extended text (2014, p.108) and argue that use of matrices allows for a "concise delivery" to the reader of what data has been analysed. Li et al (2000) concur, arguing that "a well-organized matrix could illustrate a pattern in the data better than a few paragraphs" (p.119).

Firstly, for a cross-case comparison of themes in the case studies I adapted Miles & Huberman's "case-ordered meta matrix" to: "coherently array the basic data for a major variable, across all cases" (1994, p.189). Secondly, this was also done with data sets from both strands and the wider variety of data sets that produced data, issues and themes to be compared and explored in the discussion chapter. This approach to bringing data together is also called a "joint display of data" (Creswell, 2014, p.223).

For this method I also consulted the summary tables method of presenting the findings from both the qualitative and quantitative data sets using an idea from Li et al's "parallel tracks" design which also uses a two strands design (2000 p.124). In this, it is not only the qualitative data that can be presented in these tables. Li et al (2000) describe how qualitative and quantitative data can be "crossed over" in order to be comparable. This is done by "converting some quantitative data into narratives and some qualitative data into matrices" (p.127). For example, quantitative data was transformed into "written narratives that summarized the most salient points in tables or graphs" (ibid). This allowed data from both strands to be put into matrices as a visual aid to converge the data allowing the researcher to see supporting and contradicting evidence of proficiency and ownership across the data sets. However even though I used these tables to organize and compare "cross-case" data and

themes from across the data sets, they do not speak for themselves but rather interpretations from these tables were written up in the discussion.

Before the convergence of the data from both strands I used techniques and concepts from Braun & Clarke (2006) to organise and categorise my qualitative data in a thematic analysis whilst the quantitative data was subject to descriptive and inferential statistics. The remainder of this chapter deals with a rationale for the use of this mixed methods approach; a justification of the use of the instruments illustrated in figure 3 (above); and closes with an account of the ethics of data collection. In the following chapter (chapter 4) a description of the collection of the individual data sets; analytic process for each set and findings is detailed set by set for ease of understanding for the reader.

3.1.2 Rationale of Approaches

3.1.2.1 Mixed Methods

In employing this mixed methods approach to my research, I am "integrating" quantitative and qualitative forms of data (Creswell, 2014, p. 4) because they both contribute to the same research framework: *an exploration of the proficiency and ownership of English of outer circle students for an inner circle academic context*. The overall rationale for combining quantitative and qualitative data is that this process "provides a more complete understanding of a research problem than either approach alone" (ibid p.4). In this way, mixed methods is a pragmatic approach to research, employing methods and instruments which can "work to provide the best understanding of a problem" (ibid. p.11).

In terms of an epistemology, the choice is one of pragmatism, which means that I chose methods that are useful for obtaining the information to explore the overarching theme in my research. A priority in choosing the design was to orient towards answering the research questions in the most pragmatic and effective way using the method most suitable for the respective research questions (Bryman, 2006). The basic maxim of pragmatism being "choose the combination or mixture of methods and procedures that works best for answering your research questions" (R. B. Johnson & Onuwuegbuzie, 2004, p. 17). In using mixed methods, there are multiple rationales but the overall one is to provide a "comprehensive analysis of a research problem" (Creswell & Plano Clark, 2011, p.15), what Bryman (2006) calls "completeness".

A mixed methods study also provides 'complementarity' whereby the quantitative and qualitative data sets collect "different but complementary data on the same topic" (Morse, 1991, p. 122). For example, in exploring the research theme of the study, I use a (positivistic) statistical analysis of test score data combined with (constructivist) interviews and surveys on perspectives of proficiency and ownership of English. The second research strand then extends the first using a qualitative case study approach where the experiences (constructivist) of individuals from the outer circle can be gauged in depth.

Different research questions are answered within the same study (Bryman, 2006) but they still reside within the same conceptual framework of the performance and proficiency of the outer circle student in inner circle academia. Criticisms of this approach include claims that mixed methods "have logical underpinnings rooted more in philosophy than in empirical reality" (Symonds & Gorard, 2010, p. 1) and Yin (2006) questions whether a study can be called mixed unless it integrates the various procedures in a study including research

questions and analysis. However, Creswell & Plano Clark argue that a study can be considered mixed if at the very least the study merges the results in the interpretation and discussion stage in a "convergent design" (2011, p. 79).

As alluded to above, mixed methods also provide "Illustration" (Bryman, 2006) whereby qualitative case studies in the second strand illustrate and extend quantitative and qualitative findings in the first strand. This rationale can also be expressed as "expansion" (Greene, Caracelli, & Graham, 1989) and "context" (Bryman, 2006). In this study, sole use of the particular quantitative data on proficiency and ownership of English would be "one-dimensional" and have nothing to "say" about the students' actual academic and linguistic performance in the context of the academic domain. In this way, the qualitative data enhances and "builds upon" the quantitative data (Bryman, 2006). Together, both types of data provide more *useful* findings for practitioners such as those involved with EAP and university admissions reflecting Brymans "utility" concept (ibid).

There is also the concept of "offsetting" (ibid) whereby both types of data (quantitative and qualitative) have weaknesses that can counterbalance against each other. For example, using case studies alone it is hard to generalise findings beyond the cases and the data may be without an empirical context. Quantitative data can provide powerful contextual data to support the case study. Overall, using a mixed methods approach allows me to merge the results from both strands to "develop a more complete understanding of a phenomenon" (Creswell & Plano Clark, 2011, p.77). In the case of this study, the phenomenon is the proficiency and ownership of English of the outer circle student and the implications of this for their academic performance. I now expand in more detail on the rationale for the particular methodological approaches in each strand.

3.1.2.2 Strand 1: Rationale for using Descriptive and Inferential Statistics

I am interested in how proficiency in the outer circle is expressed in the form of differences in variation in PTEA scores according to nationality. The data set provided by Pearson for 2014 allowed me to see how outer circle proficiency compares to that of countries of the expanding circle. In this way I am interested in "differences in group performance" (Bachman, 2004, p.209) in the PTEA. By running inferential statistical tests on the data, I want to know whether observed values (mean test scores in this case) from two or more groups "might all belong to the same population regardless of group or whether the observations in at least one of the groups seem to come from a different population" (Rowntree, 1991, p. 143). The findings would indicate whether a group's variation in scores are statistically significant from the others. As such, given enough data in a sample, there are statistical techniques which not only provide descriptive statistics of mean averages (for a straight comparison of group test performance) but also inferential statistics (t-tests and ANOVA) which test for whether differences in these mean averages are comparable, that is, whether they are statistically significant.

Therefore statistical tests help to "make inferences and test hypotheses about differences between different samples" (ibid p.209) helping to determine if differences in group means could be down to *real differences* in ability between groups or down to *chance* in the particular sample. Interpretation of these statistics brings an important quantitative perspective to the mixed methods research and together with interpretation of the qualitative data of the first strand gives a more rounded picture of the proficiency of the outer circle test taker.

3.1.2.3 Strand 2: Rationale for a Case Study Approach

A case study approach using four cases is used in the second strand of my research. Case studies follow one or a number of "cases" through an experience in the world. A case within a case study is "a phenomenon of some sort occurring in a bounded context" (Miles & Huberman, 1994, p.25). "Bounded" meaning that the case studies can be either tightly defined and delineated by time and activity as in an event, a school, a pupil or a class for example (Stake, 1995) or they can also be an example or instance where the boundaries between the object of study and the context it is in, are blurred (Verschuren, 2003; Yin, 2009). Taking these definitions into consideration, my cases or "subjects" (Thomas, 2011) are individual outer circle students sharing the following characteristics:

- Students of any age or gender from any 'outer circle' nation who have taken the PTEA and are able to provide an official PTEA score report.
- Entered university courses in the 2013-2014 year.
- Are studying in the UK on undergraduate, postgraduate or diploma courses.
- Are in any discipline of study.
- May or may not be accessing EAP classes.

Case studies can also be defined by type. My case study research is longitudinal and multiple:

Longitudinal

The research question on academic performance demands that the case study is longitudinal (Yin, 2009), in my case, lasting one academic year. It would be pointless to conduct a study of this nature solely over the first term when students are unlikely to have written any

assignments, nor had real time in which to identify any language or study support needs. Therefore, the study was designed to collect data from the participants at three points in the first academic year; the first point, after the first month of entry to their courses in order to gain their initial impressions of their proficiency and performance; a second point a few months into their courses to gauge the same information and to collect written evidence of performance (assignments); and a third point near the end of their first year in order to obtain more assignments and students' reflections on their experiences with the perspective of a year's study. A major strength of longitudinal case study research is in its adaptability or flexibility in design whereby one stage of the research process can inform the next (Lloyd-Jones et al 2012). For example, the outcomes of my first student tutorials influenced the content of the second student tutorials. Participants raised issues in the first sessions that were probed further in subsequent sessions.

Multiple

In using a case study approach in the second strand of my mixed methods research, I relied (principally) for my data collection on "multiple participant meanings" (Creswell, 2009, p. 8). The multiple participants provided perspectives or interpretations on linguistic and academic performance in the context of university, bounded by time in terms of their first year experiences and reflections. The term "multiple" in my case study means I have four cases which offer variety as opposed to just one case – called a "collective" case study (Stake, 1995). The significance of having multiple cases as opposed to just one case is that I can analyse data within a case and between the cases or "settings" (Baxter & Jack, 2008, p. 550). This provides an insight into the variability and similarities between cases in my study. Multiple can also refer to the fact that there are many sources (multiple sources) of data about a single case. For example, a case study participant can be given a questionnaire, interview, and observation and be assessed by someone other than themselves or the researcher – for

example a tutor. Multiple data sources are a "hallmark" of the case study approach (ibid) and "a major strength" of case studies (Yin, 2009) because they offer multiple perspectives on an issue that helps to explain and illuminate a phenomenon. One source may shed light on the research questions that is hidden by another. In my study what students say about their proficiency and ownership may be contradicted by the perspective of assignment grades, written work, self-rating, or even by themselves in other tutorials.

By taking a case study approach, I am looking for the perspectives of multiple participants (cases) and evidence - students, grades, scores and documents in order to "construct" my study and guide my data collection and analysis. In this way, I am closest to a constructivist approach which means that I "seek to understand the how and why of phenomena from a holistic, participant informed perspective" (P. A. Duff, 2008, p. 33). A constructivist approach (Baxter & Jack, 2008; Creswell, 2009; P. A. Duff, 2008), emphasises that truth is "relative" and dependent on "perspective" (Baxter & Jack, 2008). In a constructivist approach "the goal of the research is to rely as much as possible on the participants' views of the situation being studied" (Creswell, 2009, p.8). I further categorise my case study research as being a combination of "exploratory" and "descriptive" aspects (P.A. Duff, 2008). It is exploratory in the sense of being able to formulate new questions which may need to be answered in further studies and descriptive in being able to answer "what" questions, as in all of my research questions.

The data sets are collected using a variety of tools. Within both strands of the research there are surveys (questionnaires), interviews or tutorials and documents used. I now include a brief description and rationale for use of each of these tools in this mixed methods research.

3.1.3 Characteristics and Rationale of the Research Tools

3.1.3.1 Surveys

In the first research strand one survey was used. This was an online survey in the form of a questionnaire distributed to students from outer circle countries. In the second strand a survey was designed primarily as a way to recruit students to the case study.

Surveys were used because they provide an efficient way to gather data from a large sample (I initially sought responses from a potentially large number of respondents). However, it is a tool notorious for having a "low percentage of return" (Cohen et al, 2011 p.411). Cohen et al state that the larger the sample, the more closed the questions in a questionnaire because this allows the responses to be processed more rapidly (2011, p. 381). Questionnaires in this study used a mixture of closed and open questions. For example, I used closed questions such as dichotomous questions where the choice could only be one of two responses and often used them as a "funnelling device" (ibid, p.383) in the questionnaire to allow respondents to bypass questions that were not relevant to them and move to another section of the questionnaire. Dichotomous and multiple choice questions gathered data on the demographic details of respondents and their test scores.

In order to gather data on attitudes and opinions I made use of one or two rating scales (Likert) on different aspects of a focused topic. For example, in the outer circle questionnaire I used a Likert scale to gather attitudes towards admissions from the point of view of outer circle students and their qualifications in English. For this rating scale I did not provide a "middle option" that would have allowed respondents to "sit on the fence". Some may

consider this problematic but I wanted to gauge on which side of agreement or disagreement the respondents positioned themselves. In addition to these scales, some open questions included space in which to comment and expand on the answers. A variety of open-ended questions in both surveys allowed for collection of views and opinions to "enable participants to write a free account in their own terms" (ibid, p.382). The more open-ended questions in the questionnaire were dominant because I was looking for "candour" from the respondents and opinions that I could code and theme to compare with the case study tutorials. However, the drawback to this is that the questionnaires could have been seen as too lengthy and this may have stopped respondents filling in a response.

By including many open-ended questions I tried to gather as many in-depth opinions as possible but I did not oblige participants to reply to all the questions so as to prevent abandonment of the survey. In actual fact most respondents wrote extensively in the open-ended questions. This led to an awareness of weaknesses in the surveys in that they were quite lengthy and therefore could have had two effects. Firstly, completion of the questionnaire by self-selecting individuals with strongly-held views on the subject matter could have been an issue. Secondly, those students who were not of this kind may have completed the questionnaire but this may have resulted in them not completing it properly, neglecting to fill in many of the optional comments sections. There were quite possibly responses that fitted both of these categories. The surveys themselves were created and hosted in an online format, making them easy to design, disseminate and collate data from. They were constructed using the survey creation tool on the "Sitebuilder" facility at the University of Warwick and hosted on the university's postgraduate student "Eportfolio" website.

3.1.3.2 Interviews

Interview tools were used in the two strands of the research. In strand 1, Pearson country representatives were interviewed. In strand 2, the case studies involved "tutorials" which in essence were semi-structured interviews. As with questionnaires, open-ended questions in the interviews allowed the researcher to delve deeper into the attitudes and reflections of the participants with the added advantage over questionnaires for the opportunity for greater and immediate follow-up. For example, in interviews clarification of points can be immediately asked for, unlike in questionnaires in which the researcher is separated from the interviewee in both time and space.

Semi-structured interviews which were conducted neither as a casual conversation nor as an interrogation using questions asked in a strict order, allowed the "best of both worlds" (Thomas, 2011) in setting specific questions whilst allowing the interviewee the freedom to talk and reveal unexpected information not anticipated in pre-planned questions. Interview "protocols" (Creswell & Plano-Clark, 2011, p. 179) or "schedules" were used to remind the interviewer of the list of topics needing coverage in the semi-structured or "interview guide approach" (Patton, 1990, p. 7). A protocol takes the form of a list of questions and procedures to be followed (O'Loughlin, 2008; Yin, 2009) but still remains fairly open and flexible. It contains details about the context of the interview and the questions which can guide the interview, as well as instructions for the interviewer so that a standard procedure can be followed for each interviewee if desired (Yin, 2009, p. 79).

A potential weakness of this approach is that although I had a list of topics to cover there was still a danger that I missed certain data from not asking each interviewee exactly the same

questions in exactly the same order. On the other hand, as a result of their design the interviews felt more natural and the interviewees may not have opened up to me in the same way if my design had been too "structured". Other general weaknesses of interviews that should be taken into account include "bias from poorly articulated questions"; "response bias"; "inaccuracies from poor recall", and "reflexivity" – the interviewee giving what they think the interviewer wants to hear (Yin, 2009:102). These are very real pitfalls of interviews which were noticed in this study but through becoming aware of these drawbacks I could manage the interviews better.

3.1.3.3 Documentary Evidence

A document can be described as a "record of an event or a process" (Cohen et al, 2011, p.249). Exams are an event and produce certificates as results, as are written assignments and policy documents. In this sense there are various types of documentary data used in the study. In the first strand of the research I accessed the "virtual documents" of university websites and online documents such as attached and downloadable files of admissions policies regarding CSELTs or alternative English language qualifications. In the second strand of the research the four cases provided their official PTEA certificates and also provided evidence of written assignments as well as their university marking schemes for written work. The documentary data however does not "speak for itself" (Cohen et al, 2011, p.253), and instead must be interpreted. Thus, the PTEA score reports were interpreted to "say something" about the student's English language proficiency, as were the written assignments. The websites meanwhile were also subject to interpretation even though, as policy documents, they were fairly clear in their language and purpose. One problem of collecting documents such as assignments is that they may be an incomplete data set, be difficult to access or have "biased selectivity" (Yin, 2009:102). This may mean for example that when collecting students'

written assignments, examples of the students' worst work could have been deliberately withheld perhaps for reasons of face or status. The documents I did see might then have been an unrepresentative sample of the students' work – perhaps their best score on a relatively easy assignment. To counter this I asked students for examples of their best and worst written work (in terms of grades) to give me an impression of the range of their work but this of course depended on the students' cooperation and willingness to do so.

3.1.4 Case Study Instrument Design and Rationale

Within the case study I used three particular tools that are important to describe in detail as part of its design: firstly, the nature of the interviews as "tutorials"; secondly, the "self-rating" tool of proficiency; and thirdly, the "experiencing academia" topics.

3.1.4.1 Tutorials

I use the term "tutorial" to describe the interviews in the case study because the intention was to build a relationship between myself as the researcher, and the student. This was in order to build trust and to maintain some continuity with the student over their first year of study. My tutorials were semi-structured but more in line with Patton's "interview guide approach" (1990, p. 206), that is to say, structured with an interview schedule or protocol. However, I purposely conducted my tutorials as more like "guided conversations" (Yin, 2009, p. 106) because each participant was unique and had different concerns and responses so that depending on the responses of the individual participants, there was usually ample room for flexibility in the types of questions asked and the ordering of them. In later tutorials, regarding the written work, the aim was to help the student identify areas in their written performance which might have needed feedback and input from a "tutor" or third party (for

example the researcher) who could offer another opinion of their work providing the student with the opportunity to reflect on their test performance and academic performance so far. Some of the tutorials started more like formal interviews but most of them were, or became more like "chats" so that the interviewees were put at ease and could see that the time was valuable in helping them to reflect on their study experiences.

3.1.4.2 The 'self-rating' tool

A large advantage of using surveys and interviews in the case studies of my research was the chance to use a "self-rating" tool as part of the evidence of proficiency. It was designed so that students could use a CEFR-linked tool to rate their own proficiency. The purpose of this was for triangulation of evidence of proficiency - comparing what the PTEA score report "said" about their proficiency, what they self-rated as, and what their tutorials revealed about their proficiency as well as analysis of their written work. This would give a picture of the students' proficiency and their "language behavior" (Ingram & Bayliss, 2007, p.9) in academia.

The tool I used employed descriptors relating to PTEA score ranges developed from the Pearson Score Guide (2012b, p.4) (appendix 4) that can also be seen in the Pearson "concordance" table of PTEA scores with the CEFR (Pearson, 2012a, p.47). These were inspired by use of IELTS descriptors in a similar self-rating tool derived from Ingram & Bayliss's "self-rating" of language behaviour (2007). The descriptors in the concordance table were separated into an overall descriptor and four sub-skills. I used the CEFR *Structured Overview of all Scales* (Council of Europe, 2001) to fill in any "blanks" in the descriptors to produce the final tool (appendix 7). The tool was first used in the recruitment

survey (see figure 2 above), which consisted of open-ended questions and some closed questions to recruit participants to the case study. I instructed students to rate their own linguistic proficiency using this tool for the four skills and an overall score. When reading the descriptors, students were not informed as to which CEFR level the descriptors alluded to, thus, the test taker had to chose a descriptor set based on their understanding of the picture of proficiency that the descriptors "draw". The descriptor sets in the survey were simply numbered from 1-5 (students did not know what the numbers referred to although they would get a sense of a level increase from 1-5 as they read). The PTEA enabling skills (appendix 6) do not have specific descriptors so I could not construct a self-rating tool for these six sub-skills.

As mentioned above, the self-rating concept is based on a similar technique in a study by Ingram & Bayliss (2006 & 2007 p.10) who asked students to *rate* their own "language behavior" once on the course. However, unlike Ingram & Bayliss, I did not ask test takers to rate themselves with reference to how they interacted with "native speakers" because native speakers are not the only legitimate actors using English at university so I do not ask students to take into consideration *who* they are interacting with in terms of native speakers or otherwise whilst rating their language proficiency.

There were two main problems using a tool of this kind, one was the time lag between when the PTEA was taken and the point in time in which the student self-rated. Thus, natural improvement in proficiency level could account for the discrepancies between levels chosen by the student and scored in the test. This time gap could be as much as a year and a half to two years, however, the tool was primarily useful in itself, not to compare with the PTEA scores as such but to simply provide another perspective on proficiency and to compare the

CEFR level chosen with the students' comments in the tutorials regarding their proficiency at university. However, as such, the second problem with this tool is that the selection of level depends on the student's interpretation of the language in the descriptor and not least whether the student actually reads the descriptors properly or chooses them in a hurry. A student's choice of level in the tool was followed up in the first tutorials and their reasons for choosing a particular level were explored.

3.1.4.3 'Experiencing Academia' Cards

To create a "user-friendly" tool for the case studies that could encourage students to talk about all the various aspects of performance in academia, I adopted situations and tasks from instruments in previous studies (as mentioned in the literature review) that looked at many aspects of "experiencing academia" such as: "interacting with classmates"; "asking questions in seminars"; "editing texts" and so on. There was a lot of overlap of these aspects from tools in various studies but eventually I used those from Woodrow, (2006:69-70); Ferguson & White, (1994:b17-22); Banerjee (2003:417); and Byrne & Flood (2005, p. 117), a full list of the tasks and situations I finally adopted from these studies can be viewed in appendix 8.

In the tutorial, participants were given a stack of cards with these "experiencing academia" tasks or situations written on them. They were asked to read them and sort them into two piles according to whether they had experienced some or no difficulties with them. They were encouraged to comment more on their choices and expand on what they had experienced regarding problems in those areas and whether they could identify linguistic or academic issues in any of the situations. This exercise provided an effective way in to talk about any challenges and problems that the student had been experiencing at university and students were able to talk freely about them. It was an effective and innovative tool because it

avoided just reading out a long list of topics or presenting the interviewees with a long list of topics on a single sheet of paper which could have been overwhelming for them. Some studies used these topics with corresponding Likert scales asking students to rate to what extent they had problems or to rank them in order (Cheng et al, 2004; Weir et al, 2013) but I did not want to produce any statistics from the exercise as I only had four cases so the aim was instead to produce spoken data which could be transcribed and put through a thematic analysis in order to provide qualitative insights into students' perspectives on their academic performance.

I now present a description of the thematic analysis approach of the qualitative data in the study and the process of coding that was employed.

3.1.5 Thematic Analysis Approach

Qualitative data such as transcribed interviews and documentary evidence consists of large amounts of text which needs to be reduced (Li, Marquart, & Zercher, 2000, p. 119) in order to be understood and for themes to be extracted. This lends itself to thematic analysis and coding (Braun & Clarke, 2006) in order to arrive at themes that are important in the final discussion.

Coding and thematic analysis was applied to four data sets in the study and so this is described in detail here instead of repeated multiple times for the four individual sets in chapter 4. Thematic analysis was applied to:

- Strand One: a) transcriptions of spoken data from the Pearson representative interviews and b) written comments in the outer circle online survey

- Strand Two: c) written comments in the recruitment survey and d) transcriptions of case study tutorials (up to three per student)

My analytic approach to the data above largely followed the six steps in Braun & Clarke's "phases of thematic analysis" (2006, p.87).

> "Thematic analysis is a method for identifying, analysing and reporting patterns (themes) within data….Thematic analysis is widely used, but there is no clear agreement about what thematic analysis is and how you go about doing it" (Braun and Clarke, 2006, p.79)

Good thematic analysis depends on a researcher being clear and explicit about how they went about the task. The steps I took in the current study were as follows:

The first step in the analysis was to transcribe verbatim the tutorials and Pearson interviews using simple conventions (appendix 9), which constituted the first informal part of the analysis (Miles & Huberman 1994). This included the process of listening to each audio recording of the tutorials two or three times against my written transcriptions (Braun & Clarke, 2006, p.87-88). The recruitment and outer circle student surveys had already been "transcribed" by the participants themselves in the form of the online written comments in the open-ended questions. That also accounts for the spelling and punctuation errors in those quotes in the findings below as I felt it would be unethical to "correct" their written responses.

I then transferred the resultant documents into software (Nvivo) which assisted in the coding process. The first data to enter the coding process was from the recruitment survey and first case study tutorials, and later on the codes were added to by the coding of the Pearson interviews and outer circle survey. However, although sharing codes, these data sets were treated separately when it came to constructing thematic descriptions (see chapter four

below). What follows is a description of the coding process for the tutorials which also applies to the other data sets but on a smaller scale.

The second step was to begin the coding. I used coding in order to condense the data (the actual chunks of utterances of participants) and sort them into categories (called "nodes" in Nvivo). Coding is the process that allows the researcher to essentially identify, label and categorise the data eventually creating a network or "tree" of these categories or "nodes" from which themes can be interpreted. Coding is a process of identifying interesting and relevant data, extracting it, chunking it together and then ordering it and comparing it with other "chunks" of data:

> "Coding is thus a data condensation task that enables you to retrieve the most meaningful material, to assemble chunks of data that go together…" (M.B. Miles et al., 2014, p. 73)

As mentioned previously, coding is also an exercise in "data reduction" (Li et al, 2000) allowing the researcher to break down the data into usable concepts and to see themes – "manageable pieces" (ibid, p.119) to compare with other data sets. I used a deductive and inductive approach to the coding (Braun & Clarke, 2006, p.83) according to whether the initial coding was linked to the specific goals of the questions asked (deductive) or whether the codes and themes came from the data itself (emergent and inductive). In the first stage of the coding I selected chunks of conversation (data) from the interviews and tutorials that covered broad topics, which reflected the general areas I was interested in identifying in the conversations.

I often coded multiple turn-taking by interviewer and interviewee because a code or issue was sometimes embedded in an "extended exchange" (Riessman, 1993) between interviewer and

interviewee. It was important to include the surrounding content in selection of the text to be coded, what Braun & Clarke call "coding inclusively" (2006:89) thus giving the coding more context. I initially took a deductive approach to coding by assigning the chunks of data to a preconceived group or "node" in Nvivo. These nodes into which I first coded the data were formulated from the research questions that I wanted to answer. For example, I was looking for evidence of specific linguistic problems in the academic performance of the participants, therefore "specific linguistic problem" became a node and is thus an example of deductive coding or an "a priori orientating construct" (M.B. Miles et al., 2014, p. 238).

In the case of coding the tutorials, the first round of deductive coding produced the nodes as below into which I had assigned the chunks of data from the interviews. For example if a participant said "I did my first degree in English in Ghana" this would be "coded" into node viii – "previous academic experience" (see table 2 below). An example of a 'chunk' of data and its coding can be seen in appendix 10. Definitions of the codes were written and can be seen in the examples below, however, these were relatively 'loose' and could be rewritten as the coding continued.

Table 2: Initial Codes and Definitions

Code No.	Code Name	Definition
i)	Specific linguistic problems (LING)	This refers to a particular problem related to language proficiency experienced by the student. Perhaps a deficiency in grammar prevented them being able to express themselves in writing, or pronunciation impeded others from understanding them, or their lack of general vocabulary made reading very hard-going or slow.
ii)	Needing language support (LANG-ST)	This refers to the student requesting or being recommended EAP language classes from struggling with proficiency and/or the demands of the course.
iii)	Academic skills problems (AC-SKI-PROB)	This refers to particular issues with academic skills such as problems with the format of essay writing, referencing and other conventions, presentation skills or research skills problems.
iv)	Needing specific skills	This refers to the student requesting or being recommended academic skills training in response to problems in this area. For

	support/training (SKI-SUP)	example, this would include a need for information technology training which is an increasingly important academic skill.
v)	Intercultural issues encountered or noticed (INTERCUL)	This refers to cultural differences impeding performance or progress, often manifested by interactional issues such as informality in dealing with tutors, feelings of not being able to approach tutors for help, not being able to question tutors or difficulties in working with other students due to cultural differences.
vi)	Preparation for the test (TEST-PREP)	This refers to the student mentioning anything about how they prepared for the PTEA.
vii)	Difficulties with the test (TEST-DIFF)	This refers to problems with the test format or content that students felt unfairly impeded their performance during the test.
viii)	Previous academic experience (PRE-AC)	This refers to the students' previous academic experience helping or impeding their academic performance.
ix)	Previous professional experience (PRE-PROF)	This refers to any previous professional experience that contributed to how students performed on the course, helping or impeding their performance.

In step three, subsequent rounds of coding allowed me to review these nodes and allowed for inductive coding, reacting to new issues that arose in the data. New nodes were created from interesting items in the data that were not anticipated or pre-conceived. There was thus an expansion of nodes and child and parent nodes, and some of the nodes from the first round of coding (above) were split or merged further and data was reassigned to different nodes. In other words, the structure of the nodes changed and the software allowed me to arrange nodes according to the relationships between them - creating multiple "parent" and "child" nodes (appendix 11). Second and third tutorials were coded into this existing tree of nodes and led to more of this inductive coding as new concepts joined the node tree and were incorporated into the node structure. Following the steps of analysis above, I reassigned data to nodes and relabelled others, redefining the nodes by looking through the actual data under each node. Some new concepts were created from this work such as "personality" as I analysed what the data were actually about (or rather my interpretation of the utterances as data). This use of deductive and inductive coding is an example of "dialectical" research procedures (Miles et

al, 2014, p.238) that advocate a use of both "top-down" and "bottom-up" analysis (Braun & Clarke, 2006, p.83).

The fourth step involved reviewing the nodes as potential themes, reading through the nodes, and the data in a "constant comparative" method (Thomas, 2011) which meant "going through data again and again" (ibid, p.171) to see if the coded data actually "fitted" into the assigned nodes or needed to be recategorised into other nodes. This reviewing involved changing the parent nodes, and rewriting definitions of them.

After breaking up and amalgamating nodes, I eventually arrived at a structure of nodes that I thought captured the essence of my data and I could now embark on step five: reading the data in order to decide on themes. I interpreted themes from reading the data within the network of nodes but as part of the process I also found that I still had to continue to condense the data by looking at the "parent" nodes and as much as possible moving all the data in those down into child nodes so as to get more specific about exactly how the data could be labelled. This "sieving" produced the "essence" of the themes in my coding. I then went through all the nodes to again make sure that what I had coded in each one should actually belong there, or whether it needed to be re-assigned to another existing node or a new node created for it. This resulted in a changed "tree of nodes" creating six distinct branches of "parent nodes" (appendix 12) with a sub-stratum of child nodes containing most of the data in a third tier. This also resembles Goertz's "three-level framework" of concepts, (2006, p. 6) whereby the actual empirical data lies at the bottom layer of nodes but connected to the parent nodes and bigger concepts.

Looking within that hierarchy of six parent nodes, I could see how much data had been assigned to each parent and each child node. This gave me an initial first impression of how significant each node seemed to be. For example, in parent node 1 (linguistic), by looking at the "child" nodes and data within each one I saw that a lot of the data was to do with encountering different varieties of English. Similarly, in terms of the coding for particular academic skills, there were far more comments about the problems of "format" in assignment writing which seemed to pose the biggest problem for my cases. The nodes now began to help me identify the main areas of concern, or "themes", for each case. These could then be analysed and presented in terms of what concerned each case in their actual language behaviour and academic performance, and from this, thematic maps could be drawn which summarised all the themes that were interpreted as significant for each case.

The Nvivo software allowed me to systematically code or categorise the data but ultimately it was my interpretation or "reading" of the resulting system of nodes that created a theme. In this way I identified, analysed and reported "patterns" within my data sets. The themes I identified captured "something important about the data in relation to the research question" (Braun & Clarke, 2006:82). This coding process revealed the variety of perceptions of the participants on various topics including their performance, proficiency and the variables that affected these things.

In the sixth step, the emphasis was on "producing the report" involving creation of a "thematic map" (Braun & Clarke, 2006, p.89) for each participant, or data set, representing: the themes from the recruitment survey; the complete set of tutorials for each case A to D; the interviews with each Pearson interviewee; and a summary of the opinions expressed in the outer circle student survey. For each data set, interviewee or case, I constructed an individual

"thematic map". This was done in Nvivo using the idea of "models" (Bazeley & Jackson, 2013) which are essentially diagrams depicting a network or "map" of themes. A thematic map is a visual representation of the researcher's interpretation and summary of the themes that emerge from the data. In the findings chapter which follows, a thematic map is displayed for each case or interviewee after a written account of each of the themes (a thematic description) illustrated by specific data from the interviews, questionnaire or tutorials. Important here in illustrating the themes is "selection of vivid, compelling extract examples" (Braun & Clarke, 2006, p.87). These descriptions and maps are what I then use to compare with the findings from other instruments in a joint display of data from both strands. The tabulation of examples of themes across the data reveals shared or contradictory themes that are then used in the discussion and to answer the research questions.

3.2 The Ethics of Data Collection

In distributing the research instruments and approaching and managing participants I followed the ethical procedures at the University of Warwick. I make specific comments here on anonymity, procedure in interviews and the use of incentives in recruiting participants.

Anonymity

All online surveys were answered anonymously and were voluntary. Real names did not have to be given and emails were voluntarily provided at the end of the survey by the participants if they gave consent to be contacted for the interview stage. Names were not used in the presentation or discussion of the results. Participants for the initial online recruitment survey were initially contacted by Pearson from their records in the form of an email with a link to my online survey. The respondents were assigned numbers from 1-47. The participants who then agreed to the rounds of tutorials in the case study were assigned letter codes, so Case A,

Case B etc… (These do not correspond to 1, 2, 3… in the recruitment survey). University names were provided by the cases but as with the cases themselves, real names were not mentioned in the results or subsequent discussions. In the text of the thesis, the institutions are referred to simply as "Case A's institution" etc, and may refer to either universities or colleges.

Procedure in Interviews

Tutorials with the cases were essentially interviews in terms of ethics. These were recorded with the permission of the interviewee and information sheets on the research were given out and consent forms signed (appendix 13). This was similarly done with the Pearson country representatives. Participants from the initial online recruitment survey were not obliged to take any further part in the case study, and only four individuals took part in the first tutorials at the beginning of the 2013/2014 academic year. Participants were informed that they could withdraw their participation at any time without any explanation or penalty.

Incentives

No financial incentive was provided for respondents to the initial online survey, but book tokens of £20 were provided to cases for each tutorial they completed as a token of thanks for their time. The participant could refuse the book token, they were not obliged to accept it if it made them feel uncomfortable, which was the choice of Case D. If the case completed all three tutorials I then provided them with an appreciation of thanks and a £50 book token. For the online survey of outer circle students a prize draw of book tokens was offered to respondents. Students were not contacted further unless they expressed an interest and

consented in the relevant section at the end of the online survey. In the event, the two winners in the prize draw did not claim their book tokens.

Chapter 4: Data Collected, Analytical Approach and Findings

I now present the data collected, the analytical procedures employed and the findings pertaining to each research question in turn. An overview is first provided of the participants and data.

4.1 Strand 1 – Proficiency in and Ownership of English

Research Strand 1 Participants

- I accessed the admissions pages on the websites of 132 UK Universities to analyse their online documents relating to their English language requirements for international students.

- I acquired the PTEA test scores of 8067 individuals for the year 2014 broken down into 4334 individuals from 27 outer circle countries and 3733 individuals from 36 expanding circle countries in order to calculate mean test scores for the overall and four sub-skill score categories according to nationality and to test for the significance of any variation of means. I also did this for L1s within country and regional groups to see if L1 was a significant variable in mean test scores.

- Publicly-available information for mean test scores according to nationality were also accessed for the IELTS and TOEFL-iBT tests using their official websites. This was data from 2013.

- Three individuals from Pearson Education were interviewed as representatives of Pearson in three outer circle countries (Nigeria, Malaysia and India) for their views on the language proficiency, ownership of English and PTEA test performance of test takers from those countries.

- 17 individuals from outer circle countries answered an online questionnaire regarding English proficiency, test performance, university admissions and ownership of English.

The research questions addressed by the data sets are displayed at the beginning of each respective section below.

4.1.1 UK University Admissions Data

RQ1 (ADMISSIONS): What do the admissions policies of universities regarding nationality and CSELTs qualifications (as proof of English proficiency) suggest about the ownership of English and the English language proficiency of outer circle students?

4.1.1.1 Data Collection

The purpose of collecting this data was to provide evidence of the proficiency of outer circle students in respect of their ability to study at tertiary level in an inner circle context (the UK). It also provided an indication of British university admissions policies regarding the acceptance of non-inner circle English language qualifications. In November and December 2014, and August 2015, I viewed the "virtual documents" of university websites on "English language requirements", including lists of alternative qualifications as proof of English

proficiency. This was done for the 132 UK institutions listed in appendix 14 which encompasses most if not all of the HEIs in the UK. This was to investigate how flexible institutions were in accepting CSELTs as *alternatives* to the SELTs approved by the UKVI. I collated the websites of 132 UK universities from lists on The Guardian newspaper website (Guardian News and Media Limited, 2015) and Universities UK (Universities UK, 2015) to check the names and websites of the main HE institutions in the UK. All UK universities have websites which provide information to international student applicants regarding the English language requirements of the university.

4.1.1.2 Analytic Approach

This documentary data was initially categorised and analysed by putting the universities into tables representing nine regions of the UK. The rationale for this initial grouping was done simply as a convenient way for the researcher to categorise the universities in collation of websites and data. The following information was entered into the tables:

- University name
- Website address
- Indication as to whether the institution would accept a CSELT and to what degree (which specific countries and qualifications mentioned).
- Comments on the content and details of university policies on CSELTs

Where it was unclear on the websites as to whether a university did indeed accept Commonwealth CSELTs as alternatives to SELTs and other international academic English tests, the institution was emailed to enquire for confirmation. This was done in the case of 31 universities. The analysis involved re-sorting the universities into content categories to

encapsulate common policies or practices among the 132 institutions surveyed and to quantify how many and which universities fell into each category (appendix 14). The categories are explained in the findings section below.

4.1.1.3 Findings – The UK University Admissions Data

The data summarising the findings from these documents is displayed in table 3 below. The numbers in the table refer to the number of UK universities in each category.

Table 3: Results of University Admissions Survey (n=132)

Category	N	Percentage
1. University website specifically stating acceptance of CSELTs as alternatives to SELTs.	23	17.42%
2. University accepting a limited number of CSELTS and/or with some conditions (e.g. an additional interview or internal test)	33	25%
3. Acceptance of CSELTs not specifically stated on university website but confirmation of acceptance after email enquiry	20	15.15%
TOTAL ACCEPTING CSELTs	**76**	**57.57%**
4. Strictly only accepting SELT on university website.	45	34.09%
5. Not specifically stated on university website but not accepting of CSELTs after email enquiry	9	6.82%
TOTAL NOT ACCEPTING CSELTs	**54**	**40.91%**
6. Did not answer enquiry	2	1.52%

Officially, it is advised by the UKVI that a SELT is required for all international students who are not from the UKVI exempt list of MESCs (Majority English-Speaking Countries) (UK Visas and Immigration, 2016a). However, on closer inspection of the information on the websites, there were many exceptions to this "advice" illustrated in the above results. Answering "yes" to CSELTs were the following groups:

i) 23 universities (17.42%) stated that they accepted outer circle qualifications as alternatives to SELTs and even provided lists on their websites. These included universities such as

Nottingham, Nottingham Trent, Loughborough, Leicester, Sheffield, Birmingham City, Swansea (see appendix 3) and Worcester. Lists were not always complete in terms of mentioning every single Commonwealth CSELT and the prominence of certain countries that were included on the lists could sometimes differ between universities. The reasons for display of certain countries on the lists more prominently than others were not investigated (although the main or most common ones were perhaps displayed as examples as there are potentially over 60 outer circle countries and territories).

ii) 33 universities (25%) accepted CSELTs without an additional SELT or IAET but listed only a few examples of CSELTs or set caveats such as stipulating the taking of an additional interview or an internal university proficiency test. These included universities such as St Andrews, Birmingham, Warwick, Cambridge, Liverpool, Durham and Imperial. The Nigerian and Ghanaian WAEC was accepted by a large number of these institutions (n= 26).

iii) Some universities didn't list CSELTs from any countries on their website but suggested that students enquire by email to see whether their country's qualifications could be used as an alternative. 20 institutions replied positively to enquiries about acceptance of the WAEC and included universities such as Glasgow, Reading, Leeds, Greenwich, University of East London, De Montfort, Manchester, Keele, Liverpool John Moores and Bradford.

Answering "no" to CSELTs were the following groups:

iv) 45 (34.09%) of the 132 universities surveyed stated very clearly on their websites that they did not accept alternatives to SELTs. These included universities such as: Aberdeen, Edinburgh, York, SOAS, UCL, Kings, LSE and Surrey.

v) As with point iii) above, other universities did not provide clear information on their website and instructed applicants to email admissions if confirmation of a qualification's acceptance was needed. However, on following this line of enquiry outer circle CSELTs were not accepted as alternatives to a SELT by 9 institutions including St George's, Kent, Southbank and Queen Mary.

vi) Only two of the universities contacted did not respond to requests for information on CSELTs and admissions - The University of West London and Queen Margaret University.

What is revealed by the survey is that the acceptability of CSELTs for university admissions as alternatives to SELTs still depends largely on the individual decisions and policies of each HE institution. A lot of the Russell Group institutions (18 out of 24) appeared flexible regarding acceptance of CSELTs (Edinburgh, Glasgow, Cardiff, Queens, Nottingham, Birmingham, Warwick, Liverpool, Manchester, Newcastle, Leeds, Sheffield, Imperial, York, Durham, Kings, Oxford, Cambridge). In contrast, a lot of the post-1992 group (former polytechnics) seemed to be more wary of allowing alternatives to SELTs (for example West of Scotland, Westminster, Cardiff Metropolitan, Southbank, London Metropolitan, South Wales, Anglia Ruskin, Sheffield Hallam, York St John, Oxford Brookes). This could be due to the sanctions imposed on London Metropolitan University by the UKVI in 2013 (Shepherd, 2013).

Another notable trend was that a refusal or acceptance of CSELTs was often expressed in an indirect way. For example some universities stated that they accepted English-medium taught students but then only specified those from only a few countries. Others claimed that they did

accept CSELTs but then specified that the student also needed to meet the Tier 4 requirements. As the Tier 4 requirements are for the student to pass a SELT, the information on acceptance of CSELTs was contradictory or suggested that the Tier 4 requirements "trumped" the universities' own policies regarding English language requirements. Interpretation of this data is done in the discussion chapter.

4.1.2 The PTEA 2014 Data Set - Descriptive and Inferential Statistics

RQ2 (PROFICIENCY DIFFERENCE):

i) Is there a difference in the proficiency of outer circle test takers compared to expanding circle test takers as reported by PTEA and other academic test scores? (IELTS and TOEFL-iBT)

ii) Is there a difference in proficiency, as reported by PTEA test scores, according to the L1 of the test taker? (English L1 vs. Other L1)

4.1.2.1 Data Collection

The data provided by Pearson was subjected to descriptive and inferential statistical tests to provide evidence for the English proficiency of outer circle students (as reported by mean test scores). The data sets provided by Pearson included the overall PTEA score and four communicative skills scores for 8067 individuals who took the PTEA in 2014. The data was inserted into SPSS software[4], which was used to organise and manipulate the data sets to perform statistical tests. Data was provided for a maximum of 500 test takers per country as stipulated by Pearson but no reason was given for this limitation. This data was a *sample* of students who took the PTEA test in 2014. In terms of outer circle and expanding circle

[4] a software package used for statistical analysis acquired by IBM in 2009

categorisation of these test takers there were 4334 test takers from 27 ‘outer circle’ countries and 3733 test takers from 36 ‘expanding circle’ countries broken down into the following nationalities as in tables 4 and 5 below:

Table 4: Outer Circle Country Data (n=4334)

		Frequency	Percent
Valid	Bangladesh	500	11.5
	Botswana	6	.1
	Cameroon	52	1.2
	Cyprus	3	.1
	Fiji	3	.1
	Gambia	10	.2
	Ghana	73	1.7
	Hong Kong	195	4.5
	India	500	11.5
	Kenya	257	5.9
	Lesotho	2	.0
	Malawi	8	.2
	Malaysia	500	11.5
	Mauritius	55	1.3
	Nigeria	500	11.5
	Pakistan	500	11.5
	Philippines	276	6.4
	Seychelles	1	.0
	Sierra Leone	14	.3
	Singapore	80	1.8
	South Africa	393	9.1
	Sri Lanka	197	4.5
	Swaziland	2	.0
	Tanzania	37	.9
	Uganda	32	.7
	Zambia	21	.5
	Zimbabwe	117	2.7
	Total	4334	100.0

Table 5: Expanding Circle Country Data (n=3733)

		Frequency	Percent
Valid	Argentina	14	.4
	Austria	4	.1
	Brazil	500	13.4

Bulgaria	26	.7
China	500	13.4
Colombia	64	1.7
Denmark	5	.1
Egypt	36	1.0
Finland	10	.3
France	110	2.9
Germany	49	1.3
Greece	24	.6
Indonesia	40	1.1
Iran	85	2.3
Israel	112	3.0
Italy	88	2.4
Japan	318	8.5
Kazakhstan	72	1.9
Korea, South	192	5.1
Libyan Arab Jamahiri	49	1.3
Mexico	89	2.4
Morocco	36	1.0
Norway	17	.5
Poland	51	1.4
Portugal	27	.7
Romania	24	.6
Russian Federation	155	4.2
Saudi Arabia	232	6.2
Spain	40	1.1
Sweden	26	.7
Switzerland	7	.2
Taiwan	38	1.0
Thailand	76	2.0
Turkey	500	13.4
United Arab Emirates	57	1.5
Viet Nam	60	1.6
Total	3733	100.0

Categorisation Issues

In the preparation of the data sets, some countries still posed some theoretical problems as to how they could be categorised, either as an outer or expanding circle country (as discussed in the literature review on problematising Kachru's circles model). However, despite these

concerns, because of the historical nature of Kachru's model and its associated legacies in education and national life in my conceptualisation of the outer circle I was content to include all the above countries in table 3 as outer circle.

4.1.2.2 Analytic Approach

The first step was to formulate a null hypothesis for the statistical tests for which I used the literature as a theoretical base. In the literature on outer circle Englishes, test takers from the outer circle applying to study abroad could, in theory, already be at a high level given their ownership of English, education in the medium of English, and the range and depth of use of English in their countries (Graddol, 2006; Higgins, 2003; Kachru, 1998; Modiano, 1999; Widdowson, 1994; Zheng & Wei, 2014). The expanding circle test performance according to country is an unknown but it may be expected to be on average lower than outer circle countries given the EFL status of English in many of these countries.

Although I can theorise that there will be a difference in test performance between the outer and expanding circles, statistical analyses usually work on the basis of a null hypothesis which exists for the statistical tests to disprove or reject. Thus I worked on the basis of the null hypothesis as follows:

"There is no statistically significant difference in the proficiency of test takers from the outer and expanding circles as expressed by mean PTEA scores"

After securing the data and the hypotheses, I ran descriptive and inferential statistics. The descriptive statistics allowed me to see what the breakdown of each sample was in terms of nationality, gender, L1, age and the mean scores and standard deviation of their overall PTEA score and the four communicative skill scores for listening, reading, speaking and writing.

The initial descriptive statistics would also reveal the groupings that could be made for the inferential tests, for example, a breakdown of mean scores for individual countries or regions allowed the researcher to see whether there were any further distinctions that could be made in test performance both within and between the outer and expanding circles.

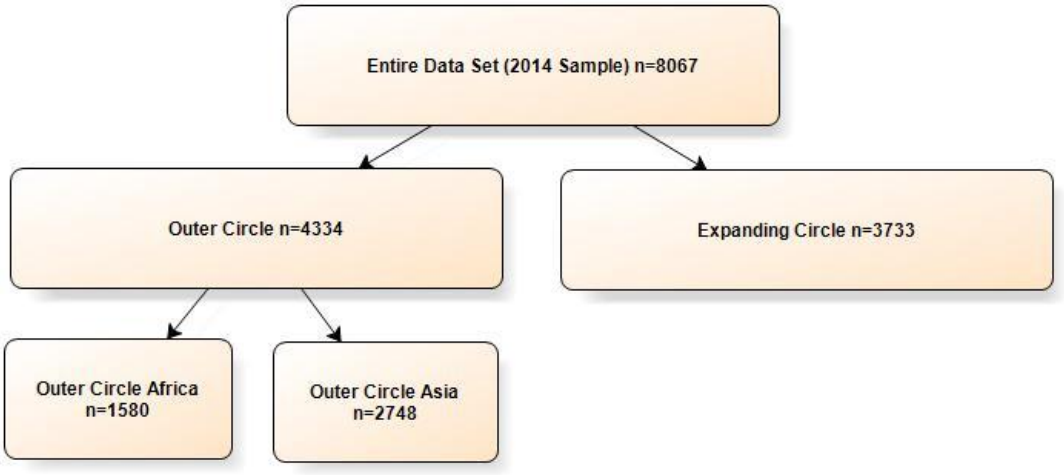
The inferential statistics comprised t-tests and ANOVA (analysis of variance) to determine if there were any statistically significant differences in the variance of mean PTEA test scores between groups. The t-test compared the variance in means of two groups while the ANOVA performed the same function but applied to three or more groups. In the case of using ANOVA, an additional test - the "post-hoc Tukey" test was used to identify where any statistically significant difference lay (between which particular groups) (Cohen et al, 2011, p.647-648).

4.1.2.3 Findings - Descriptive Statistics for Main Groups

The descriptive statistics helped to illustrate the "normality" of the sample in terms of the gender balance, age ranges, proportion of test takers to country and range of L1s and this data can be seen in the descriptive profiles of each group in tables 6-22 below.

I first present the PTEA mean scores and other descriptive statistics for the entire data set of 8067 individuals and each sub-group in figure 4 below, which I have defined for my later t-test and ANOVA analysis. On the following pages I present the descriptive profiles for the groups and sub-groups illustrated in Figure 4:

Figure 4: Hierarchy of the Data Groupings for t-test and ANOVA



Note: Cyprus (n=3) and Fiji (n=3) are present in the outer circle data set but not the Africa or Asia sub-sets so that explains the n=6 discrepancy between those two tiers.

As can be seen in figure 4, the entire data set includes the expanding and outer circle groups. The outer circle group (OC) includes the subgroups OC Africa and OC Asia. The original aim was to compare the variation in mean PTEA scores of the outer and expanding circles. The outer circle was further split into Africa and Asia because differences in mean scores were noticed in the descriptive statistics for individual countries making up these two regions (table 24 below) which led to a need to compare the variation in mean scores between these two groups and the expanding circle group in an ANOVA (see below).

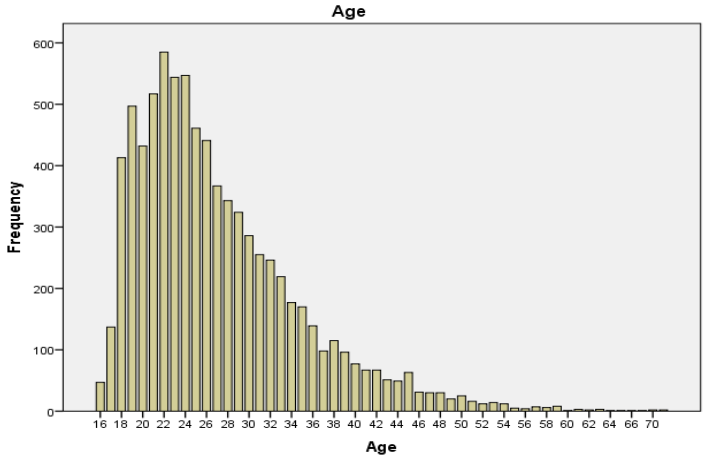
Table 6 illustrates the first set of descriptive statistics from the data provided by Pearson: the number of participants in each group; the proportions of first-time or repeater test-takers; male and female test takers, and statistics for age including mean and range.

Table 6 Test Taking Status, Gender and Age Descriptive Stats for all Main Groups

		Entire Set	Expanding Circle	Outer Circle	Outer Circle Africa	Outer Circle Asia
N		8067	3733	4334	1580	2748
Test Taking Status	First Time	n=6138 (76.1%)	n=3007 (80.6%)	n=3131 (72.2%)	n=1306 (82.7%)	n=1820 (66.2%)
	Repeater	n=1929 (23.9%)	n=726 (19.4%)	n=1203 (27.8%)	n=274 (17.3%)	n=928 (33.8%)
Gender	Female	n=3624 (44.9%)	n=1913 (51.2%)	n=1711 (39.5%)	n=742 (47%)	n=965 (35.1%)
	Male	n=4443 (55.1%)	n=1820 (48.8%)	n=2623 (60.5%)	n=838 (53%)	n=1783 (64.9%)
Age	Mean	27.04	26.80	27.25	28.92	26.29
	SE	.086	.125	.118	.242	.121
	SD	7.720	7.620	7.800	9.610	6.335
	Range	55: (16-71)	52: (16-68)	55: (16-71)	55: (16-71)	54: (16-70)

Most people in each group are first-time test takers of the PTEA. OC Asia has the largest proportion of repeater test takers at a third (33.8%) and OC Africa the smallest (17.3%). Most groups have more male than female test takers with the exception of the expanding circle group. The mean age of a test taker for the entire data set is around 27 years of age. OC Africa has the highest mean age of test takers at nearly 29. The age distribution for the entire data set is depicted below (figure 5) with most test-takers being at the younger end of the scale, as might be expected for university students seeking to study abroad. This pattern was repeated for all the groups in table 6 above.

Figure 5: Age Range & Distribution for the Entire Data Set (n=8067)



In the tables below (8-22), each of the groups is displayed in terms of their mean PTEA scores with standard deviation, followed by data illustrating the composition of the groups in terms of number of test takers per country and L1 speakers per named language as well as the country in which the test was taken. Throughout the presentation of the descriptive and inferential findings, I use the following colour coding (table 7) to illustrate the "fit" of the GSE scale to the equivalent levels in the CEFR and the B2 threshold for university set by the UKVI. The same colour coding is used and referred to throughout the thesis (also used for the case study self-rating data).

Table 7: PTEA Scores (GSE) and CEFR Levels Colour Key (Pearson, 2012a, p.47)

CEFR	C1+	B2	B1+	B1	A2	A1
Score (GSE)	76+	59-75	51-58	43-58	30-42	10-29

Explanations for the data follow each table starting with the entire data set:

Table 8: Mean PTEA Scores for the Entire Data Set

Entire 2014 Sample (n=8067)					
PTEA Score Category	Overall	Listening	Reading	Speaking	Writing
Score (GSE)	55.15	55.26	56.95	53.10	57.80
Stan Dev.	16.945	17.728	17.242	20.438	16.185

Table 8 indicates that the mean scores for the entire data set are between 51 and 58 on the GSE for each score category (Overall, Listening, Reading, Speaking & Writing). This is aligned to B1+ on the CEFR (yellow). The standard deviation is similar across the categories (16-17) apart from the largest standard deviation, which is for the speaking score (20.438). This pattern of standard deviation can be detected throughout the data sets for other groups and individual countries (see tables 11, 14, 17, 20, 23 and 24 below).

Table 9: Entire Sample (n=8067) Composition of Nationality and L1s

Entire 2014 Sample Breakdown by Country		Entire 2014 Sample Breakdown by L1	
Country	% (n)	Language (L1)	% (n)
1. China	6.2% (n=500)	1. English	25.8% (n=2082)
2. Turkey	6.2% (n=500)	2. Mandarin	7.8% (n=630)
3. Brazil	6.2% (n=500)	3. Portuguese	6.4% (n=550)
4. Nigeria	6.2% (n=500)	4. Turkish	6.1% (n=490)
5. Malaysia	6.2% (n=500)	5. Bengali	4.8% (n=391)
6. Pakistan	6.2% (n=500)	6. Urdu	4.8% (n=389)
7. India	6.2% (n=500)	7. Arabic	4.7% (n=377)
8. Bangladesh	6.2% (n=500)	8. Japanese	3.7% (n=295)
9. South Africa	4.9% (n=393)	9. Cantonese	2.8% (n=229)
10. Japan	3.9% (n=318)	10. Russian	2.3% (n=182)
11. Philippines	3.4% (n=276)	11. Korean	2.2% (n=175)
12. Kenya	3.2% (n=257)	12. Spanish	2.1% (n=170)
13. Saudi Arabia	2.9% (n=232)	13. French	2.1% (n=167)
14. Sri Lanka	2.4% (n=197)	14. Tagalog	2.0% (n=163)
15. Hong Kong	2.4% (n=195)	15. Malay	2.0% (n=159)
16. South Korea	2.4% (n=192)	16. Telugu	1.6% (n=167)
17. Russia	1.9% (n=155)	17. Hindi	1.5% (n=118)
18. Zimbabwe	1.5% (n=117);	18. Tamil	1.4% (n=112)
19. Israel	1.4% (n=112)	19. Punjabi	1.3% (n=106)
20. France	1.4% (n=110)	20. Hebrew	1.0% (n=83)
21. Mexico	1.1% (n=89)	21. Sinhalese	1.0% (n=78)
Other countries (n=43)	<1% each	Other L1s (n=54+)	<1% each

The breakdown by country (table 9) gives an idea of which countries provided most test-takers in the sample. The top eight countries have a maximum of 500 individuals allowed for the samples provided by Pearson (no reason given) so the real total number of test takers from each country for 2014 could be much higher. Despite an upper limit of 500 participants for eight of the countries, having unequal numbers of test takers from the other countries is realistic as the number of participants per country do seem somewhat in proportion to their populations, for example South Africa with a population of 54,956,000 (The World Bank,

2015) provided data for 393 test takers while the Seychelles with a population of 92,000 provided data for only 1 test taker. The breakdown by L1 matches to some extent the country data. For example Portuguese is high on the list reflecting Portuguese-speaking Brazil (n=500), and the high number of Mandarin speakers (n=500) reflects the large sample from China. A quarter of the entire data set speaks English as an L1 (25.8%).

Table 10: Entire Sample (n=8067) Test Centre Locations

Entire 2014 Sample Breakdown of Test Centre Location	
Country	% (n)
1.UK	41.4% (n=3339)
2. Turkey	6.2% (n=499)
3. Brazil	5.5% (n=441)
4. Malaysia	5.4% (n=438)
5. India	5.4% (n=434)
6. Nigeria	4.4% (n=356)
7. China	3.4% (n=278)
8. Kenya	3.2% (n=259)
9. Japan	2.9% (n=236)
10. Hong Kong	2.3% (n=184)
11. USA	2.1% (n=172)
12. South Africa	1.9% (n=157)
13. Australia	1.7% (n=141)
14. Singapore	1.5% (n=118)
15. South Korea	1.3% (n=105)
16. Saudi Arabia	1.3% (n=103)
17. UAE	1.1% (n=85)
Other countries (n=30)	Under 1% each

In table 10, just over 40% of the test takers took the test in the UK. However, it is not possible to know the length of stay of each test taker in the UK before they took the PTEA.

The next tables (11-13) display data for the expanding circle:

Table 11: Mean PTEA Scores for the Expanding Circle

Expanding Circle (n=3733)					
PTEA Score Category	Overall	Listening	Reading	Speaking	Writing
Score (GSE)	52.61	51.81	56.05	49.45	55.09
Stan Dev.	17.305	17.961	17.562	20.674	15.987

In comparison to the entire data set, the mean scores of the expanding circle set are slightly lower in all score categories but still lie between 51 and 58 GSE (B1+), with the exception of the speaking category (49.95 GSE).

Table 12: Expanding Circle (n=3733) Composition of Nationality and L1s

EC Breakdown by Country		EC Breakdown by L1	
Country	% (n)	Language (L1)	% (n)
1. China	13.4% (n=500)	1. Portuguese	13.7% (n=513)
2. Turkey	13.4% (n=500)	2. Turkish	13.1% (n=488)
3. Brazil	13.4% (n=500)	3. Mandarin	12.2% (n=457)
4. Japan	8.5% (n=318)	4. Arabic	10.0% (n=372)
5. Saudi Arabia	6.2% (n=232)	5. English	8.0% (n=298)
6. South Korea	5.1% (n=192)	6. Japanese	7.9% (n=294)
7. Russia	4.2% (n=155)	7. Russian	4.9% (n=182)
8. Israel	3.0% (n=112)	8. Korean	4.7% (n=175)
9. France	2.9% (n=110)	9. Spanish	4.6% (n=170)
10. Mexico	2.4% (n=89)	10. French	3.1% (n=114)
11. Italy	2.4% (n=88)	11. Hebrew	2.2% (n=83)
12. Iran	2.3% (n=85)	12. Italian	1.8% (n=68)
13. Thailand	2.0% (n=76)	13. Thai	1.8% (n=67)
Other countries (n=23)	<2.0% each	14.Vietnamese	1.4% (n=52)
		15. German	1.2% (n=43)
		16. Persian	1.0% (n=36)
		Other L1s (n=46+)	<1.0% each

English as an L1 drops down to only 8% of the sample (from 25% in the entire data set) suggesting that many L1 English test takers lie in the outer circle group. The EC L1 proportions more or less reflect the proportions of nationality.

Table 13: Expanding Circle (n=3733) Test Centre Locations

EC Breakdown of Test Centre Location	
Country	% (n)
1. UK	33.6% (n=1254)
2. Turkey	13.3% (n=497)
3. Brazil	11.8% (n=439)
4. China	6.9% (n=257)
5. Japan	6.3% (n=235)
6. South Korea	2.8% (n=105)
7. France	2.0% (n=73)
8. Israel	1.7% (n=65)
9. Australia	1.4% (n=54)
Others countries (n=54)	<1.0% each

A third (33.6%) of the total expanding circle test takers (n=3733) took the PTEA in the UK.

Tables 14-16 are for the outer circle:

Table 14: Mean PTEA Scores for the Outer Circle

Outer Circle (n=4334)					
PTEA Score Category	Overall	Listening	Reading	Speaking	Writing
Score (GSE)	57.34	58.23	57.72	56.24	60.14
Stan Dev.	16.317	16.975	16.926	19.701	15.991

Compared to the expanding circle, the outer circle PTEA mean scores have risen towards GSE 59, with higher scores for writing just within the CEFR B2 band (60.14). The outer circle test takers as a group have higher mean PTEA scores than EC test takers in all score categories. However, regarding the CEFR levels, the outer circle and expanding circle groups both display a B1+ level in the overall, listening and reading score categories.

Table 15: Outer Circle (n=4334) Composition of Nationality and L1s

OC Breakdown by Country		OC Breakdown by L1	
Country	% (n)	Language (L1)	% (n)
1. Bangladesh	11.5% (n=500)	1. English	41.2% (n=1784)
2. India	11.5% (n=500)	2. Bengali	8.0% (n=391)
3. Malaysia	11.5% (n=500)	3. Urdu	8.0% (n=388)
4. Nigeria	11.5% (n=500)	4. Cantonese	4.5% (n=194)
5. Pakistan	11.5% (n=500)	5. Mandarin	4.0% (n=173)
6. South Africa	9.1% (n=393)	6. Tagalog	3.7% (n=161)
7. Philippines	6.4% (n=276)	7. Malay	3.6% (n=158)
8. Kenya	5.9% (n=257)	8. Telugu	2.9% (n=126)
9. Sri Lanka	4.5% (n=197)	9. Hindi	2.7% (n=116)
10. Hong Kong	4.5% (n=195)	10. Tamil	2.5% (n=109)
11. Zimbabwe	2.7% (n=117)	11. Punjabi	2.4% (n=105)
12. Singapore	1.8% (n=80)	Others L1s (n=45+)	<1.0% each
13. Ghana	1.7% (n=73)		
14. Mauritius	1.3% (n=55)		
15. Cameroon	1.2% (n=52)		
Others countries (n=12)	<1.0% each		

Asian countries comprise much of the outer circle sample (Bangladesh, India, Pakistan and Malaysia), with Nigeria the other “large” test-taking nation. The proportion of test takers with

English as their L1 has risen to 41.2% of the sample compared to 25% for the entire data set and 8% for the expanding circle. All the other main L1s are Asian languages.

Table 16: Outer Circle (n=4334) Test Centre Locations

OC Breakdown of Test Centre Location		
Country	% (n)	
1. UK	48.1%	(n=2085)
2. Malaysia	9.8%	(n=424)
3. India	9.7%	(n=421)
4. Nigeria	8.2%	(n=354)
5. Kenya	5.9%	(n=257)
6. Hong Kong	3.7%	(n=161)
7. South Africa	3.6%	(n=156)
8. Singapore	2.2%	(n=95)
9. Australia	2.0%	(n=87)
10. Philippines	1.2%	(n=52)
11. USA	1.2%	(n=52)
12. Zimbabwe	1.2%	(n=51)
13. Ghana	1.2%	(n=50)
Other countries (n=16)	<1.0% each	

Almost half of the outer circle sample (48.1%) took the PTEA in the UK.

Tables 17-19 are for OC Africa:

Table 17: Mean PTEA Scores for OC Africa

Outer Circle Africa (n=1580)					
PTEA Score Category	Overall	Listening	Reading	Speaking	Writing
Score (GSE)	63.95	65.43	64.83	61.36	67.76
Stan Dev.	16.573	17.065	17.195	21.062	15.381

Once the data is separated into OC Africa and OC Asia, some dramatic changes become evident in the data. Viewing the data for OC Africa (table 17), all the score categories show a mean score in the B2 CEFR band and all are above 60 on the GSE. Writing is the category with the highest mean score at 67.76 GSE.

Table 18: OC Africa (n=1580) Composition of Nationality and L1s

OC Africa Participant Breakdown by Country		OC Africa Participant Breakdown by L1	
Country	% (n)	Language (L1)	% (n)
1. Nigeria	31.6% (n=500)	1. English	73.2% (n=1156)
2. South Africa	24.9% (n=393)	2. Afrikaans	4.8% (n=76)
3. Kenya	16.3% (n=257)	3.Yoruba	3.7% (n=58)
4. Zimbabwe	7.4% (n=117)	4. French	3.4% (n=53)
5. Ghana	4.6% (n=73)	5. Igbo	2.3% (n=36)
6. Mauritius	3.5% (n=55)	6. Swahili	2.2% (n=34)
7. Cameroon	3.3% (n=52)	7. Hausa	1.6% (n=26)
8. Tanzania	2.3% (n=37)	8. Twi	1.0% (n=16)
9. Uganda	2.0% (n=32)	Other L1s (n=24+)	<1.0% each)
10. Zambia	1.3% (n=21)		
11. Sierra Leone	0.9% (n=14)		
12. Gambia	0.6% (n=10)		
13. Malawi	0.5% (n=8)		
14. Botswana	0.4% (n=6)		
15. Lesotho	0.1% (n=2)		
16. Swaziland	0.1% (n=2)		
17. Seychelles	0.1% (n=1)		

In table 18, over 72% of the sample comes from just three countries (Nigeria, South Africa and Kenya), and nearly three-quarters of the sample (73.2%) have English as an L1.

Table 19: OC Africa (n=1580) Test Centre Locations

OC Africa Breakdown of Test Centre Location	
Country	% (n)
1. UK	42.4% (n=670)
2. Nigeria	22.3% (n=353)
3. Kenya	16.3% (n=257)
4. South Africa	9.9% (n=156)
5. Zimbabwe	3.2% (n=51)
6. Ghana	3.2% (n=50)
Other countries (n=21)	<1.0% each

A large proportion of African test takers (42.4%) took the test in the UK.

Table 20: Mean PTEA Scores for OC Asia

	Outer Circle Asia (n=2748)				
PTEA Category	Overall	Listening	Reading	Speaking	Writing
Score (GSE)	53.55	54.10	53.66	53.32	55.78
Stan Dev.	14.910	15.491	15.364	18.244	14.654

In comparison to OC Africa, OC Asia displays lower PTEA mean scores at between 51 and 58 on the GSE (B1+) and English as an L1 drops to 22.8% of the sample (still a lot higher than in the expanding circle sample).

Table 21: OC Asia (n=2748) Composition of Nationality and L1s

OC Asia Participant Breakdown by Country		OC Asia Participant Breakdown by L1	
Country	% (n)	Language (L1)	% (n)
1. Bangladesh	18.2% (n=500)	1. English	22.8% (n=626)
2. India	18.2% (n=500)	2. Bengali	14.2% (n=391)
3. Malaysia	18.2% (n=500)	3. Urdu	14.1% (n=388)
4. Pakistan	18.2% (n=500)	4. Cantonese	7.1% (n=194)
5. Philippines	10.0% (n=276)	5. Mandarin	6.3% (n=173)
6. Sri Lanka	7.2% (n=197)	6. Tagalog	5.9% (n=161)
7. Hong Kong	7.1% (n=195)	7. Malay	5.7% (n=157)
8. Singapore	2.9% (n=80)	8. Telugu	4.6% (n=126)
		9. Hindi	4.2% (n=115)
		10. Tamil	4.0% (n=109)
		11. Punjabi	3.7% (n=103)
		12. Sinhalese	2.8% (n=77)
		Other L1s (n=15+)	<1.0% each)

Non-English L1s make up a large proportion of L1s in the sample, representing their importance in certain countries. For example Bengali in Bangladesh, Urdu in Pakistan, Cantonese in Hong Kong, Tagalog in the Philippines, Malay in Malaysia and a variety of L1s in India and Sri Lanka (Telugu, Hindi, Punjabi, Tamil and Sinhalese).

Table 22: OC Asia (n=2748) Test Centre Locations

OC Asia Breakdown of Test Centre Location	
Country	% (n)
1.UK	51.3% (n=1409)
2. Malaysia	15.3% (n=420)
3. India	15.1% (n=416)
4. Hong Kong	5.9% (n=161)
5. Singapore	3.5% (n=95)
6. Australia	2.9% (n=81)
7. Philippines	1.9% (n=52)
8. USA	1.8% (n=50)
Other countries (n=10)	<1.0% each)

A large proportion of OC Asian test takers took the PTEA in the UK (51.3%) with Malaysia and India also significant locations for test taking (15% each).

4.1.2.4 Findings - Descriptive Statistics for Individual Countries

Having reported the above patterns of mean scores within the larger groups, I explored any variations in mean scores according to particular countries. I display below the mean scores and standard deviation for, firstly, the expanding circle countries, and then those of the outer circle. As mentioned above in table 7, colour coding also indicates CEFR levels.

Expanding Circle Countries and Mean Scores

These mean score statistics per country (table 23) serve as indicators of the proficiency of test takers from individual outer circle countries. Countries in the entire data set for which there was data but only for very small samples of 10 or less are not displayed below (Finland n=10, Switzerland n=7, Denmark n=5 and Austria n=4). I list the countries in order from lowest to highest mean scores in the overall score category on the PTEA.

Table 23: Expanding Circle Countries Mean Scores (from lowest to highest overall score)

		Overall	Listening	Reading	Speaking	Writing
1. Libya	GSE	40.24	39.75	41.55	39.38	42.00
(n=49)	SD	11.466	12.946	12.218	14.527	11.828
2. Turkey	GSE	42.24	40.62	46.70	38.95	45.59
(n=500)	SD	15.197	15.979	15.134	18.389	13.468
3. Thailand	GSE	43.22	42.82	45.16	39.05	47.14
(n=76)	SD	14.983	15.615	14.837	17.097	13.832
4. Vietnam	GSE	43.92	42.60	47.73	37.33	49.03
(n=60)	SD	17.220	18.422	16.091	18.479	16.021

5. China (n=500)	GSE	46.43	44.87	49.63	41.01	50.31
	SD	12.666	12.974	13.181	14.841	11.586
6. Japan (n=318)	GSE	46.95	45.99	51.24	40.94	51.29
	SD	14.178	14.502	14.634	16.502	13.285
7. Saudi Arabia (n=232)	GSE	46.96	47.33	48.13	45.61	49.63
	SD	17.562	17.991	18.163	20.593	17.498
8. South Korea (n=192)	GSE	48.27	47.27	51.42	44.62	51.19
	SD	15.887	16.673	15.292	19.216	15.091
		Overall	Listening	Reading	Speaking	Writing
9. Iran (n=85)	GSE	50.49	50.27	51.89	51.58	51.66
	SD	16.322	17.513	16.186	20.647	14.672
10. Taiwan (n=38)	GSE	51.11	49.92	53.92	48.76	52.82
	SD	17.864	18.117	17.886	22.509	15.300
11. Kazakhstan (n=72)	GSE	54.13	53.68	57.33	52.39	57.71
	SD	15.603	15.221	16.247	21.302	15.487
12. Egypt (n=36)	GSE	55.05	56.25	56.00	56.52	56.25
	SD	17.267	18.670	17.332	21.895	19.014
13. UAE (n=57)	GSE	55.12	56.02	55.39	57.53	55.81
	SD	14.119	14.727	15.352	18.539	13.925
		Overall	Listening	Reading	Speaking	Writing
14. Indonesia (n=40)	GSE	57.83	57.03	59.95	55.50	61.35
	SD	18.999	19.981	18.740	22.606	16.443
15. Brazil (n=500)	GSE	57.99	56.87	63.55	53.53	59.65
	SD	14.045	14.698	14.427	17.680	13.484
16. Spain (n=40)	GSE	58.75	57.80	62.47	54.90	61.02
	SD	15.773	16.273	15.585	18.470	15.227
17. Colombia (n=64)	GSE	59.31	57.78	62.18	55.15	61.89
	SD	15.56	16.00	15.69	20.54	13.78
18. Norway (n=17)	GSE	59.35	58.41	61.29	61.17	59.82
	SD	15.410	18.960	14.810	16.838	15.021
19. Italy (n=88)	GSE	59.61	59.07	63.38	55.77	61.35
	SD	15.305	15.501	15.871	18.398	14.827
		Overall	Listening	Reading	Speaking	Writing
20. Bulgaria (n=26)	GSE	60.26	62.50	60.61	55.38	63.23
	SD	16.165	15.373	19.405	21.639	13.467

21. Morocco (n=36)	GSE	61.91	61.61	63.41	62.25	62.94
	SD	21.071	20.995	20.990	23.890	18.864
22. Poland (n=51)	GSE	61.92	62.73	65.57	58.65	64.14
	SD	15.068	16.350	16.465	19.549	16.156
		Overall	Listening	Reading	Speaking	Writing
23. Romania (n=24)	GSE	62.70	64.33	65.50	59.83	64.83
	SD	14.057	15.533	15.494	14.872	15.479
24. Israel (n=112)	GSE	64.34	63.34	67.34	66.81	62.31
	SD	18.151	18.172	18.680	20.283	16.931
		Overall	Listening	Reading	Speaking	Writing
25. Mexico (n=89)	GSE	64.54	63.62	67.85	64.65	64.60
	SD	15.591	16.480	15.451	20.373	14.061
26. France (n=110)	GSE	64.69	65.03	68.29	60.18	67.08
	SD	13.692	14.062	14.287	16.402	12.870
27. Russia (n=155)	GSE	65.01	64.77	68.06	63.45	66.58
	SD	17.044	17.194	16.982	20.240	16.202
28. Portugal (n=27)	GSE	67.37	65.96	70.51	68.44	65.77
	SD	15.711	16.479	15.950	18.004	15.363
		Overall	Listening	Reading	Speaking	Writing
29. Sweden (n=26)	GSE	68.73	68.07	73.53	67.34	69.30
	SD	13.739	14.265	12.050	20.368	12.591
30. Argentina (n=14)	GSE	69.64	70.42	72.00	68.42	70.50
	SD	17.18	17.85	15.76	20.34	15.53
31. Germany (n=49)	GSE	69.69	70.47	70.71	69.98	69.73
	SD	17.684	17.395	17.531	21.347	15.528
32. Greece (n=24)	GSE	70.75	69.33	74.45	74.45	69.95
	SD	10.665	12.024	10.623	15.404	12.249

All of the countries scoring under B1+ overall from the 2014 sample (50 or less on the GSE) are from the Middle East, East & SE Asia, with only Iran, South Korea and Japan reaching B1+ in some of the score categories. Among the countries scoring at B1+ overall, Spain, Brazil and Indonesia score into the B2 CEFR band in some score categories. Next listed are the higher mean scoring countries at B2 overall, although some of the sample sizes for these

higher scoring countries are quite low (Norway, Romania, Bulgaria, Portugal, Sweden, Greece and Argentina) suggesting that a larger sample of at least 30 (Field, 2013, p. 172) from these countries would have been preferable.

Summary of Expanding Circle Country-specific Data

The mean scores for expanding circle countries revealed that the expanding circle contains roughly equal numbers of countries that are in the lower (below B1+), medium (B1+) and higher (B2+) proficiency bands in terms of the CEFR. In other words, some countries display mean scores that are deemed much lower than university-level proficiency (below B1+), some are nearly at or around the threshold (B1+), while others are at the threshold (B2).

Outer Circle Countries and Mean Scores

I now display the individual country mean scores for outer circle countries. Those countries which I had data for, but had samples that were too small for inclusion here were Malawi (n=8); Botswana (n=6); Lesotho (n=2); Swaziland (n=2) and the Seychelles (n=1). For the inferential statistics some of these countries were incorporated into the group 'Southern Africa' along with Zimbabwe (n=117) as they also share features of regional English with Zimbabwe (Kamwangamalu, 2009). Fiji (n=3) and Cyprus (n=3) were left out of these two data sets altogether (OC Africa and OC Asia) because Cyprus is in Europe and Fiji is in The Pacific. Below are the mean scores for the outer circle countries in order of lowest to highest overall scores.

Table 24: Outer Circle Countries Mean Scores (from lowest to highest overall score)

		Overall	Listening	Reading	Speaking	Writing
1. Bangladesh	GSE	45.55	45.91	45.27	46.40	47.63
(n=500)	SD	10.182	10.376	10.841	13.920	9.985
2. Sri Lanka	GSE	48.49	48.73	48.59	49.41	49.94
(n=197)	SD	12.071	12.335	12.435	16.877	11.734
3. Pakistan	GSE	48.55	48.93	48.04	51.25	49.80
(n=500)	SD	12.203	12.554	12.738	16.780	11.491
4. Hong Kong	GSE	50.79	51.18	51.69	45.92	54.55
(n=195)	SD	15.555	15.907	15.560	19.123	14.109
5. Cameroon	GSE	53.10	54.52	53.60	47.52	59.06
(n=52)	SD	11.630	12.818	12.884	13.452	13.233
		Overall	Listening	Reading	Speaking	Writing
6. Nigeria	GSE	55.48	57.54	56.34	49.85	61.38
(n=500)	SD	14.323	15.858	14.692	17.053	14.576
7. India	GSE	56.54	57.02	55.95	58.88	57.69
(n=500)	SD	14.069	14.648	14.412	18.630	13.318
8. Philippines	GSE	58.82	59.68	59.57	57.66	62.21
(n=276)	SD	15.541	16.130	15.552	19.320	15.315
9. Uganda	GSE	59.46	59.96	59.34	62.18	61.84
(n=32)	SD	15.862	16.841	16.747	18.431	14.542
10. Sierra Leone	GSE	59.57	61.07	58.85	61.78	62.21
(n=14)	SD	20.413	21.427	20.717	22.330	20.737
11. Malaysia	GSE	59.91	60.79	60.92	55.45	63.55
(n=500)	SD	14.486	15.603	14.611	17.949	14.317
12. Tanzania	GSE	61.10	62.62	59.40	64.67	63.05
(n=37)	SD	15.194	15.576	15.305	20.366	12.174
		Overall	Listening	Reading	Speaking	Writing
13. Ghana	GSE	61.46	62.98	61.02	60.41	65.83
(n=73)	SD	13.168	15.283	14.055	16.265	14.138
14. Zimbabwe	GSE	62.20	64.23	63.17	57.46	67.84
(n=117)	SD	16.517	17.097	16.282	22.48	15.386
15. Mauritius	GSE	63.50	66.30	63.16	58.03	68.61
(n=52)	SD	14.603	14.770	15.447	21.241	13.699
16. Zambia	GSE	65.57	66.14	66.80	66.09	69.52

(n=21)	SD	15.64	16.44	15.81	21.54	14.97
17. Kenya	GSE	66.68	67.99	66.96	66.76	69.45
(n=257)	SD	14.868	15.310	15.895	19.402	13.569
18. South Africa	GSE	76.12	76.76	78.25	75.31	77.29
(n=393)	SD	13.723	13.863	13.750	18.088	13.239
19. Singapore	GSE	77.49	78.50	78.45	74.33	78.60
(n=80)	SD	11.295	11.630	11.428	14.127	11.314

Summary of Outer Circle Mean Scores

Among outer circle African countries, Nigeria and Cameroon produce the lowest mean scores and are B1+ overall with a B2 level only in writing. There is a difference in proficiency between Anglophone Cameroon and Francophone Cameroon, which is explored in table 39 below. Apart from Nigeria and Cameroon, all the other outer circle African countries were at B2 level overall and for most of the skill scores, although two of the countries were small in sample size (Sierra Leone and Zambia). Most of these states are multilingual and use English as a lingua franca unlike many other states in OC Asia and the EC group, which would tend to have a non-English L1 as a lingua franca. South Africa is the most proficient African country according to an overall mean score at 76.12 GSE (C1). Within OC Asia, Singapore with 77.49 GSE (C1) followed by Malaysia with 59.91 GSE (B2) have the highest mean scores overall whilst the Philippines (58.82 GSE) and India (56.54 GSE) are at a B1+ level overall, very close to the UKVI threshold for university study in the UK. Hong Kong displays a low level of English at 50.79 GSE (B1 overall) but at B1+ in three of the communicative skill categories, followed by the south Asian countries of Pakistan, Sri Lanka and Bangladesh performing at B1 in every score category except speaking in Pakistan (B1+). The most proficient nations among the outer circle countries are South Africa and Singapore.

This descriptive comparison between outer and expanding circle countries and a breakdown of mean scores according to countries revealed was that it was useful to make a further comparison between OC Africa, OC Asia and the expanding circle in an inferential statistics test. Before inferential statistics were explored, tests for normality were carried out on the data.

4.1.2.5 Findings – Inferential Statistics Tests for Normality

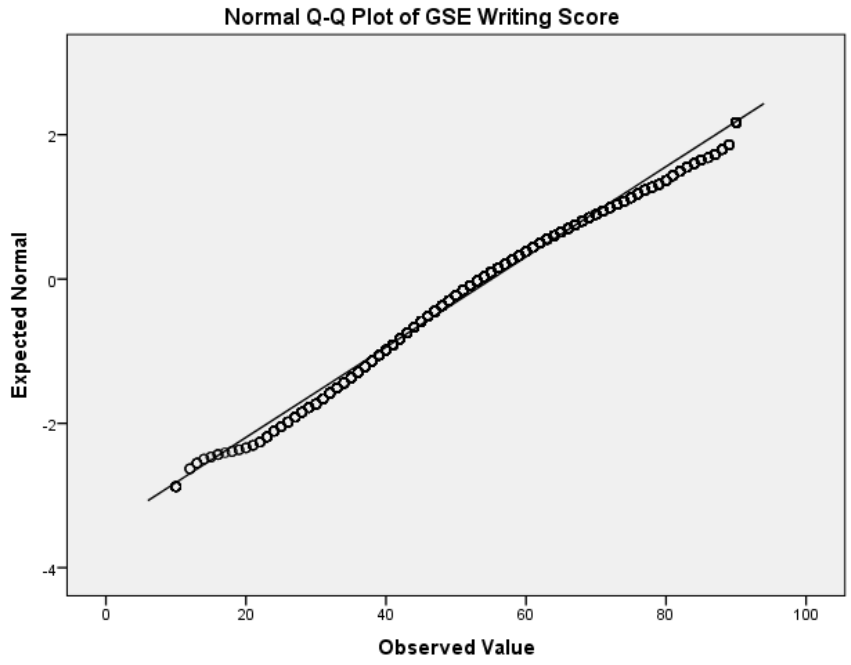
Before embarking on inferential tests of the data to see whether the variation of means between the groups are statistically significant, tests of normality were required to see whether the data for each group was normally distributed and therefore reliable for use in inferential tests (Pallant, 2010, pp. 59-62; Tokunaga, 2015, pp. 140-141). The normality tests included the following:

- Trimmed means
- Test of Normality (Komonorov-Smirnov & Shapiro-Wilk)
- Normal Q-Q plot
- Histograms
- Skewness & Kurtosis
- Detrended Normal Q-Q plot
- Box plot (identifying outliers)

I do not report a detailed analysis of each of the tests for each of the skills, but summarise in the interests of brevity. Various tests for normal distribution of the data indicated that the data for the expanding circle and outer circle data sets was generally normally distributed. The closeness of trimmed means to the sample means indicated that any outliers in the sets (also

indicated in box-plots) did not have a disproportionate effect on the sample means. This is true of both the outer and expanding circles data. The Komolgorov-Smirnov and Shapiro-Wilks tests (both at p=0.00 for both groupings) indicated that the samples violated the assumption of normality, however this is common in larger samples whereby these tests for normality are "notoriously affected by large samples in which small deviations from normality yield significant results" (Field, 2013 p.877). Another method used to indicate that the sample was generally normally distributed was the Q-Q plot with an example for the expanding circle writing score category displayed below (figure 6) displaying a pattern which was repeated for the other skills scores and for the outer circle data. The graph compares two lines superimposed on each other; the straight line represents the expected value from a normal distribution of the data and the line of circles represents the observed value of each score (Pallant, 2010 p.63).

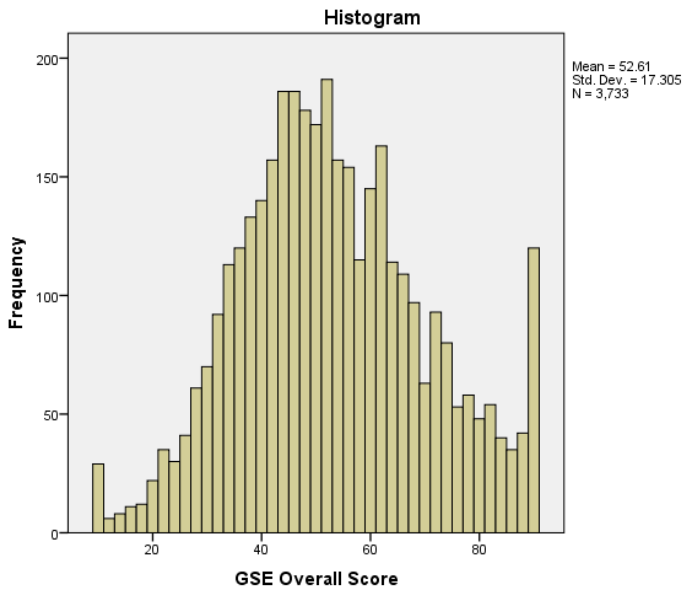
Figure 6: Tests of Normality - Example of Q-Q Plot EC Writing Scores



A histogram was also used to determine normality. A normal distribution using a histogram is represented by an imaginary line drawn over the top of the bars in a histogram to form a bell-curve. The curve in the histograms for my data was generally a bell-curve for each skill score

with a clustering of scores at the high end of the x axis on the graph (see figure 7 below). The clustering can be explained because a score of 90 on the GSE is a "false" limit to the concept of proficiency expressed on a continuous but limited scale such as the GSE (from 10 but stopping at 90). This is because, in reality, it may be difficult to claim that everyone who is advanced "ends" their proficiency at the score of 90. In other words, many test takers might score higher than 90 in terms of real proficiency (some "advanced" individuals are more "advanced" than others) and so there was a "bunching" of scores at the higher end of the histogram. This is known as a "ceiling effect". A more realistic or "normal" distribution of the scores can be imagined to be more evenly distributed beyond the arbitrary score of 90 and therefore be better represented by a smooth curve downwards beyond 90 on the scale. Without an extension of the curve beyond a score of 90, this explains the "spike" at the extreme right of all the histograms for the 5 respective skill areas. An example graph for the expanding circle overall score category is given below and illustrates this phenomenon. This was a typical histogram representing the distribution of scores for all the score categories for both the expanding and outer circle data:

Figure 7: Tests of Normality - Example of Histogram EC Overall Scores



Skewness and Kurtosis were two further indicators used to determine the normal distribution of data. The degree of skewness and kurtosis affects the reliability of statistics or inferences that can be made from the statistics (Cohen et al, 2011, p.612). The skewness or kurtosis in my data did not exceed 2.00 or –2.00 for any of the score categories and so the data can be said to be reasonably normally distributed (Bachman, 2004, p.74). The results of the tests of normality for the outer circle data set were almost identical to the expanding circle above. Having been judged reasonably normally distributed, inferential testing was then carried out on the data sets. The first inferential statistics (below) compared the variation in means of the various circles of English groups.

4.1.2.6 Findings - Inferential Statistics I – OC, EC, OC Africa and OC Asia

Before reporting on the inferential statistics (t-test and ANOVA) it is important to explain the calculation and interpretation of effect sizes.

Accompanying the results of the t-tests and ANOVA are indications of the effect size of the statistics (only relevant where differences in variation are found to be statistically significant). According to Field (2013), an effect size is "An objective and (usually) standardized measure of the magnitude of an observed effect" (p.874). This standardisation allows us to "compare effect sizes across different studies that have measured different variables, or have used different scales of measurement" (ibid, p.79). Put another way, an effect size can be equated to an indication of the size of the "influence" of the independent variable on the dependent variable (Cramer & Howitt, 2004, p. 102). In this study this is the influence of the variable "nationality" or "L1" on the mean PTEA mean score. The effect size for the inferential statistics (t-tests) has been calculated using two methods, Cohen's *d* and Pearson's *r*. This is because "when group sizes are very discrepant, *r* can be quite biased

compared to *d*" (Field, 2013:83). In my data sets group sizes do sometimes show large differences in size, suggesting that Cohen's *d* may be a preferable calculation to use (McGrath & Meyer, 2006). In the ANOVA tests a different method for calculating effect size is used (eta squared) and was calculated manually by dividing the sum of squares between groups by the total sum of squares. Figures for these were produced in the test outputs from SPSS.

Cohen's *d* and Pearson's *r* were calculated using the effect size calculators on the website of the University of Colorado, Colorado Springs (Becker, 2016) and following the procedures reported in Field (2013), Tokunaga (2015) and Pallant (2010). Table 25 indicates what the numbers in the inferential results below actually mean in terms of effect sizes (small, medium or large).

Table 25: Summary of Effect Size Types and Meanings

Effect Size Calculation	Small Effect	Medium Effect	Large Effect
Cohen's d[5]	0.2	0.5	0.8
Pearson's r[6]	.10 (the effect accounts for 1% of the variance)	.30 (the effect accounts for 9% of the variance)	.50 (the effect accounts for 25% of the variance)
Eta squared[7]	.01	.06	.14

In the results of the t-tests and ANOVAs an additional indicator of the effect size is displayed next to the effect size statistic as follows: (L) large, (M) medium, (S) small. I now display the results of the inferential statistical tests.

[5] Tokunaga (2015, p. 408)

[6] Field (2013, p.82)

[7] Pallant (2010, p.262-263), Cohen et al (2011, p.216) Field (2013, p.472)

4.1.2.6.1 t-test Results for OC and EC Groups.

The results for the t-test comparing the mean scores of the outer circle test takers with the expanding circle test takers (tables 11 and 14) in the five score categories are presented below. In these tables, the initials SE stand for "standard error of measurement" and the initials SD stand for "standard deviation".

Table 26: t-test Results for Expanding Circle (EC) vs. Outer Circle (OC)

t-test Results					
	Overall	Listening	Reading	Speaking	Writing
Levene's test Sig	.009*	.081	.040*	.010*	.014*
Sig (two-tailed)	.000	.000	.000	.000	.000
t	-12.566	-16.467	-4.344	-15.031	-14.138
df	7731.452	8065	7795.151	7762.724	7889.401
Effect Size (Cohen's *d* and Pearson's *r*)					
Cohen's *d*	0.285 (S)	0.366 (S)	0.098 (S)	0.341 (S)	0.318 (S)
r	0.141 (S)	0.180 (S)	0.049 (S)	0.168 (S)	0.157 (S)

*** Equal variances not assumed**

The t-test above suggests that there are statistically significant differences between the two groups with small effect sizes in all score categories. In other words, the outer circle does display a statistically higher proficiency in English than the expanding circle group with a small effect size and the outer circle is closer to the university proficiency threshold set by the UKVI (B2) than the expanding circle group (tables 11 and 14).

However, the descriptive statistics broken down by country (tables 23 and 24 above) indicated that there may be further differences in mean scores, specifically for outer circle Asia and Africa and the expanding circle. Therefore, it was considered relevant to compare the three group means for the expanding circle, outer circle Africa and outer circle Asia in an ANOVA test to discover if the differences indicated in the descriptive statistics were statistically significant.

4.1.2.6.2 ANOVA Results – OC Africa, OC Asia and EC

Before performing the ANOVA for a comparison of the three groups I conducted normality tests on the new outer circle sub-groups of "outer circle Africa" and "outer circle Asia". The patterns of normality generally mirrored those for the outer and expanding circle groups. The table below displays the results of the ANOVA between the expanding circle, outer circle Africa and outer circle Asia groups, taking account of the mean test scores in tables 11, 17 and 20:

Table 27: ANOVA for Expanding Circle vs. OC Africa vs. OC Asia

ANOVA					
	Overall	Listening	Reading	Speaking	Writing
Levene's test Sig	.000*	.000*	.000*	.000*	.000*
Sig (p)	.000	.000	.000	.000	.000
	Post Hoc (Tukey)				
OC Africa mean difference (sig p=0.05)	EC 11.341 (.000) OC Asia 10.394 (.000)	EC 13.617 (.000) OC Asia 11.333 (.000)	EC 8.777 (.000) OC Asia 11.165 (.000)	EC 11.905 (.000) OC Asia 11.034 (.000)	EC 12.664 (.000) OC Asia 11.980 (.000)
OC Asia mean difference (sig p=0.05)	OC Africa -10.394 (.000)	EC 2.284 (.000) OC Africa -11.333 (.000)	EC -2.388 (.000) OC Africa -11.165 (.000)	EC 3.871 (.000) OC Africa -11.034 (.000)	OC Africa -11.980 (.000)

	Effect Size (eta squared)				
Eta Sq.	0.066 (M)	0.083 (M)	0.054 (S-M)	0.046 (S-M)	0.092 (M)

*** Equal variances not assumed**

All the tests violated the homogeneity of variance assumption, which is indicated by a statistic of less than .05 on Levene's statistic (represented by the letter 'F'). The table above shows a statistic of .000 for each score category. Therefore I needed to consult statistics for the "Robust Test of Equality of Means". These statistics (Welch & Brown-Forsythe) are two "corrected versions" of the F ratio and they indicated a statistical significance of .000 ($p<.05$) for all the score categories. The post-hoc Tukey tests indicated between which groups this statistical significance lay (also indicated in table 27).

The general finding from the inferential statistical tests is that there is a statistically significant difference in the overall means and all four sub-skill score means between OC Africa on the one hand, and the expanding circle and OC Asia on the other. OC Africa displays higher mean scores in all score categories than the other two groups, which is statistically significant with medium effect sizes in most categories with the smallest effect size in speaking. The Tukey test also shows that OC Asia has a statistically significant higher mean score in listening and speaking than the expanding circle group but a statistically significant lower mean for reading.

4.1.2.7 Findings - Inferential Statistics II – L1 as a Variable

Further t-tests and ANOVAs were conducted within regional and country groups in order to see whether having English as an L1 was a further factor in test performance within the groups and individual countries. The results of these tests are described next. The first

descriptive statistics and first set of t-test and ANOVA (above) only identified proficiency differences according to the country of origin of the test takers. These tests explored proficiency differences according to another variable, the L1 of the test taker, another dimension to the nationality identity of the test taker.

Independent-samples t-tests and One-Way ANOVAs were conducted to compare the mean PTEA overall score and the four communicative skills scores of listening, reading, speaking and writing between test takers with English as a stated first language (English L1) and test takers with a stated first language other than English (Other L1). This was performed on a variety of groups as listed below. The tests were used to see whether any identified differences in the descriptive mean scores were statistically significant or non-significant. ANOVAs were conducted where there were enough participants (n>30) to compare more than one language group of non-English L1 test takers with a group of L1 English test takers from that country (for example in the case of Malaysia comparing "L1 English" Malaysian test takers with "L1 Malay" Malaysians and "L1 Chinese" Malaysians).

There are five larger group comparisons and eleven smaller regional and country groupings to encompass all of the outer circle countries in the given sample from 2014:

Large Comparisons (all t-tests)

1. The entire data set
2. The expanding circle
3. The outer circle
4. Outer circle Africa
5. Outer circle Asia

Outer Circle country & Regional Groups (t-tests unless marked otherwise)

1. East Africa (Kenya, Tanzania, Uganda)
2. Ghana
3. Nigeria (ANOVA)
4. Cameroon (descriptive only)

5. South Africa
6. Southern Africa (Zimbabwe, Zambia, Malawi, Botswana, Lesotho, Swaziland)
7. Bangladesh
8. India (ANOVA)
9. Malaysia (ANOVA)
10. Pakistan
11. Philippines
12. Sri Lanka (ANOVA)

Two outer circle nations or territories are absent from the list above – Singapore and Hong Kong. Table 28 below indicates the numbers and percentages of L1 English test takers within individual outer circle countries. Singapore does not have enough non-English L1 test takers to compare against the vast majority of the sample (88.8%) who were English L1 test takers, which is a unique situation among the outer circle Asian countries. Conversely, in the case of Hong Kong there were not enough English L1 participants to compare with Cantonese L1 test takers, and the territory has the lowest percentage of English L1 test takers in the outer circle data set (6.2%). What is also noticeable about the table is that the outer circle African countries all have higher percentages of their test takers stating English as an L1 than any OC Asian nation apart from Singapore.

Table 28: Test takers in OC countries and regions with L1 English (2014 sample)

OC Country/Region	Test Takers in 2014 sample (n)	Test Takers in sample stating English as L1 (n)	Percentage of test takers in the sample with English as L1	Percentage of test takers in the sample with Other L1
1. Singapore	80	71	88.8	11.2
2. Uganda	32	28	87.5	12.5
3. Kenya	257	218	84.8	15.2
4. East Africa+	326	266	81.6	18.4
5. South Africa	393	317	80.7	19.3
6. Zimbabwe	117	91	77.8	22.2
7. Southern Africa++	156	121	77.6	22.4
8. Nigeria	500	351	70.2	28.8
9. Tanzania	37	20	54.1	45.9
10. Cameroon	52	28	53.8	46.2
11. Ghana	73	37	50.7	49.3
12. Mauritius	55	20	36.4	63.6
13. Philippines	276	90	32.6	67.4

14. Malaysia	500	152	30.4	69.6
15. Sri Lanka	197	47	23.9	76.1
16. Bangladesh	500	111	22.2	77.8
17. Pakistan	500	80	16	84
18. India	500	63	12.6	87.4
19. Hong Kong	195	12	6.2	93.8

+ (Kenya, Uganda, Tanzania combined)

++ (Swaziland, Lesotho, Botswana, Zambia, Zimbabwe, Malawi combined)

Regarding the expanding circle (table 29 below) it can be seen that there were not enough English L1 test takers from any expanding circle country for a t-test or ANOVA to be meaningfully performed (n>30).

Table 29: Proportion of test takers in EC Countries with English as L1

EC Country/Region	Test Takers in 2014 sample (n)	Test Takers stating English as L1 (n)	% Of test takers in the sample with English as L1
1. Portugal	27	8	29.6
2. Greece	24	7	29.1
3. Argentina	14	4	28.6
4. Poland	51	14	27.4
5. Bulgaria	26	7	26.9
6. Sweden	27	7	26.9
7. Colombia	64	14	21.9
8. Egypt	36	7	19.4
9. Iran	85	15	17.6
10. Italy	88	15	17.0
11. Russia	155	26	16.8
12. France	110	18	16.4
13. Germany	49	8	16.3
14. Spain	40	6	15
15. UAE	57	8	14.0
16. Morocco	36	5	13.9
17. Israel	112	12	10.7
18. Thailand	76	8	10.5
19. Mexico	89	9	10.1
20. Kazakhstan	72	7	9.7
21. Romania	24	2	8.3
22. Japan	318	19	6.0
23. S. Korea	192	10	5.2
24. Saudi Arabia	232	12	5.2
25. Turkey	500	13	2.6

26. Brazil	500	10	2.0
27. China	500	8	1.6
28. Norway	16	0	0

The high percentage of English L1s in many of the small samples could explain the high mean PTEA scores on the test for these nations in the 2014 sample (see previous section table 23).

t-test and ANOVA results

Note: The tables below contain the descriptive statistics for the data sets and the inferential results (t-tests or ANOVA) as well as effect sizes where needed. They are colour coded according to the following key, with B1 level in white:

Table 30: CEFR levels (colours) relating to PTEA GSE score ranges

CEFR Level	C1+	B2	B1+	B1
PTE Academic Score Range	76+	59-75	51-58	43-58

Larger Comparisons (all t-tests)

1. The Complete Data Set (outer and expanding circle combined)

Table 31: Entire Data Set t-test for English L1 vs. Other L1

PTEA Score Categories & Mean Scores (Entire 2014 Sample n=8067)					
Group	Overall	Listening	Reading	Speaking	Writing
English L1	63.6	64.8	64.3	62.2	66.3
(n=2082)	SE 0.3	SE 0.3	SE 0.3	SE 0.4	SE 0.3
25.8%	SD 17.2	SD 17.7	SD 17.8	SD 21.0	SD 16.3
Other L1	52.2	51.9	54.3	49.2	54.8
(n=5985)	SE 0.2	SE 0.2	SE 0.2	SE 0.2	SE 0.1
74.2%	SD 15.8	SD 16.4	SD 16.2	SD 19.2	SD 15.0
		t-test Results			
Levene's test Sig	0.00*	0.00*	0.00*	0.00*	0.00*

Sig (two-tailed)	0.00	0.00	0.00	0.00	0.00
t	26.608	29.089	22.574	23.523	28.378
df	3379.512	3417.562	3351.034	3361.735	3383.013
Effect Size (Cohen's *d* and Pearson's *r*)					
Cohen's *d*	0.915 (L)	0.995 (L)	0.779 (ML)	0.811 (L)	0.975 (L)
r	0.416 (M)	0.445 (M)	0.363 (M)	0.375 (M)	0.438 (M)

***Equal Variances not assumed.**

According to descriptive statistics there is a noticeable difference in the mean PTEA scores (in all score categories) of L1 English test takers compared to Other L1 test takers in the entire data set (both outer and expanding circle). Those with English as an L1 score more than 10 GSE points higher than those with Other L1s in every score category. Looking at the CEFR levels, L1 English test takers display levels at B2 whilst those with other L1s are B1+. The t-test revealed statistically significant differences between the two groups in the overall and all four skill score categories. Using Cohen's *d*, the magnitude of the effect is large in all score categories, with the exception of a medium effect in speaking. When using Pearson's *r* as an alternative measure of magnitude, the effect size is medium for all score categories.

2. The Expanding Circle (EC)

Table 32: Expanding Circle t-test for English L1 vs. Other L1

PTEA Score Categories & Scores (Expanding Circle n=3733)					
Group	Overall	Listening	Reading	Speaking	Writing
English L1	62.2	62.5	64.7	60.7	64.0
(n=298)	SE 1.0	SE 1.0	SE 1.0	SE 1.2	SE 0.9
8%	SD 18.1	SD 18.6	SD 18.4	SD 22.2	SD 16.8
Other L1	51.7	50.8	55.3	48.4	54.3
(n=3435)	SE 0.2	SE 0.3	SE 0.2	SE 0.3	SE 0.2
92%	SD 16.9	SD 17.5	SD 17.2	SD 20.2	SD 15.6

t-test Results					
Levene's test Sig	0.013*	0.010*	0.025*	0.00*	0.008*
Sig (two-tailed)	0.00	0.00	0.00	0.00	0.00
t	9.671	10.364	8.488	9.210	9.654
df	343.846	344.292	343.726	341.057	343.144
Effect Size (Cohen's *d* and Pearson's *r*)					
Cohen's *d*	1.043 (L)	1.117 (L)	0.915 (L)	0.997 (L)	1.042 (L)
r	0.462 (M)	0.487 (M)	0.416 (M)	0.446 (M)	0.462 (M)

***Equal Variances not assumed**

According to descriptive statistics there is a noticeable difference in the mean PTEA scores (in all score categories) of L1 English test takers compared to Other L1 test takers in the expanding circle data set (with a difference of more than 10 GSE points in each score category). Looking at CEFR levels, L1 English test takers display a high proficiency level at B2 whilst those with other L1s are at B1+ and B1 levels. The t-test reveals statistically significant differences between the two groups in the overall score and all four skills score categories with large effect sizes in all categories using Cohen's *d* but nearing a medium effect size in all categories using the *r* calculation.

3. The Outer Circle (OC)

Table 33: Outer Circle t-test for English L1 vs. Other L1

PTEA Score Categories & Scores (Outer Circle n=4334)					
Group	Overall	Listening	Reading	Speaking	Writing
English L1	63.8	65.1	64.3	62.5	66.7
(n=1784)	SE 0.4	SE 0.4	SE 0.4	SE 0.4	SE 0.3
41.2%	SD 17.0	SD 17.5	SD 17.7	SD 20.8	SD 16.2

Other L1 (n=2550) 58.8%	52.7 SE 0.2 SD 14.0	53.3 SE 0.2 SD 14.7	53.1 SE 0.2 SD 14.6	51.8 SE 0.3 SD 17.5	55.5 SE 0.2 SD 14.0
t-test Results					
Levene's test Sig	0.00*	0.00*	0.00*	0.00*	0.00*
Sig (two-tailed)	0.00	0.00	0.00	0.00	0.00
t	22.530	23.355	22.003	17.608	23.724
df	3353.981	3404.869	3345.428	3404.112	3479.344
Effect Size (Cohen's *d* and Pearson's *r*)					
Cohen's *d*	0.778 (ML)	0.800 (L)	0.760 (ML)	0.603 (M)	0.804 (L)
r	0.362 (M)	0.371 (M)	0.355 (M)	0.288 (SM)	0.373 (M)

***Equal Variances not assumed**

According to descriptive statistics, there is a noticeable difference in the mean PTEA scores (in all score categories) of L1 English test takers compared to Other L1 test takers in the outer circle data set, with a difference of 10 GSE points or more in each score category. Looking at CEFR levels L1, English test takers have a high mean score at B2 whilst those with other L1s are at B1+ level in all score categories. The t-test reveals statistically significant differences between the two groups in the overall score and all four skills score categories with large effect sizes in the categories of overall, listening, reading and writing using Cohen's *d* but modest effect sizes using *r*.

4. Outer Circle Africa

Table 34: OC Africa t-test for English L1 vs. Other L1

PTEA Score Categories & Scores (Outer Circle Africa n=1580)					
Group	Overall	Listening	Reading	Speaking	Writing

English L1	65.6	67.2	66.4	63.5	69.0
(n=1156)	SE 0.4	SE 0.4	SE 0.5	SE 0.6	SE 0.4
73.2%	SD 16.4	SD 16.7	SD 17.1	SD 21.1	SD 15.1
Other L1	59.4	60.6	60.5	55.3	64.1
(n=424)	SE 0.7	SE 0.8	SE 0.8	SE 0.9	SE 0.7
26.8%	SD 16.1	SD 16.8	SD 16.4	SD 19.6	SD 15.4
		t-test Results			
Levene's test Sig	.165	.852	.067	.001*	.701
Sig (two-tailed)	0.00	0.00	0.00	0.00	0.00
t	6.665	6.916	6.124	7.237	5.740
df	1578	1578	1578	803.083	1578
		Effect Size (Cohen's *d* and Pearson's *r*)			
Cohen's *d*	0.335 (S)	0.348 (S)	0.308 (S)	0.510 (M)	0.288 (S)
r	0.165 (S)	0.171 (S)	0.152 (S)	0.247 (S)	0.143 (S)

***Equal variances not assumed**

A large majority of African outer circle test takers (73.2%) state English as their L1. According to descriptive statistics there is a noticeable difference in the mean PTEA scores (in all score categories) of L1 English test takers compared to Other L1 test takers with a 6-8 GSE score difference in each score category. However, regarding the CEFR levels, Other L1 test takers display the same level of proficiency as English L1 test-takers (at B2) in all score categories apart from speaking (B1+). This suggests that OC African test-takers perform quite well on the test regardless of whether or not they have English as an L1 (B2 level). The t-test suggests there are statistically significant differences between the two groups in every score category but with small effect sizes – the only exception being for speaking (medium).

5. Outer Circle Asia

Table 35: OC Asia t-test for English L1 vs. Other L1

PTEA Score Categories & Scores (Outer Circle Asia n=2748)					
Group	Overall	Listening	Reading	Speaking	Writing
English L1 (n=626) 22.8%	60.6 SE 0.7 SD 17.7	61.5 SE 0.7 SD 18.2	60.5 SE 0.7 SD 18.2	60.5 SE 0.8 SD 20.2	62.5 SE 0.6 SD 17.2
Other L1 (n=2122) 77.2%	51.4 SE 0.2 SD 13.2	51.9 SE 0.3 SD 13.8	51.6 SE 0.2 SD 13.7	51.1 SE 0.3 SD 17.0	53.7 SE 0.2 SD 13.1
t-test Results					
Levene's test Sig	0.00*	0.00*	0.00*	0.00*	0.00*
Sig (two-tailed)	0.00	0.00	0.00	0.00	0.00
t	11.975	12.200	11.340	10.539	11.682
df	841.070	849.141	845.252	902.004	848.888
Effect Size (Cohen's *d* and Pearson's *r*)					
Cohen's *d*	0.825 (L)	0.837 (L)	0.780 (ML)	0.701 (ML)	0.801 (L)
r	0.381 (M)	0.386 (M)	0.363 (M)	0.331 (M)	0.372 (M)

*** Equal Variances not assumed**

According to the descriptive statistics there is a noticeable difference in the mean PTEA scores of OC Asian L1 English-speakers compared to Other L1s, with a difference of around 10 on the GSE in all score categories. L1 English test takers display a high-level of proficiency (at B2) whilst those with other L1s are just into the B1+ level in all score categories. For OC Asia, the results indicate that there are noticeable differences between the mean scores of English L1s and non-English L1s in all five score categories. The t-test suggests there are statistically significant differences between the two groups in every score category with large or medium-large effect sizes indicated in all categories when using Cohen's *d*, but medium effect sizes with *r*.

Outer Circle Country and Regional Groups (all t-tests unless stated)

6. East Africa (Kenya, Uganda, Tanzania)

Table 36: East Africa t-test for English L1 vs. Other L1

PTEA Score Categories & Scores (East Africa n=326)					
Group	Overall	Listening	Reading	Speaking	Writing
English L1 (n=266) 81.6%	66.3 SE 0.9 SD 15.2	67.6 SE 0.9 SD 15.6	66.5 SE 0.9 SD 16.2	67.0 SE 1.2 SD 19.6	68.7 SE 0.8 SD 13.6
Other L1s (n=60) 18.4%	61.0 SE 1.8 SD 14.3	61.8 SE 1.9 SD 14.9	60.2 SE 1.9 SD 15.0	61.8 SE 2.2 SD 17.6	64.3 SE 1.7 SD 13.8
t-test Results					
Levene's test	.495	.841	.392	.396	.909
Sig					
Sig (two-tailed)	.014	.009	.007	.064	.025
t	2.462	2.616	2.718	1.860	2.254
df	324	324	324	324	324
Effect Size (Cohen's *d* and Pearson's *r*)					
Cohen's *d*	0.273 (S)	0.290 (S)	0.302 (S)	n/a	0.250 (S)
r	0.135 (S)	0.143 (S)	0.149 (S)	n/a	0.124 (S)

A large majority of East African test takers (81.6%) state English as their L1. According to descriptive statistics, there is a difference in the mean PTEA scores (in all score categories) of L1 English test takers compared to Other L1 test takers with a GSE point difference of 4-6 in each score category. However, looking at the CEFR levels, Other L1 test takers display the same high level of performance as English L1 test takers (all above GSE 60 in all score categories at B2). This suggests that East African test takers of PTEA perform quite well on the test regardless of whether or not they have English as an L1. The t-test results suggest that there are statistically significant differences between the two groups in every score category except speaking, but that the magnitude of the effect is small using both Cohen's *d* and *r*.

7. Ghana

Table 37: Ghana t-test for English L1 vs. Other L1

PTEA Score Categories & Scores (Ghana n=73)					
Group	Overall	Listening	Reading	Speaking	Writing
English L1 (n=37) 50.7%	62.3 SE 2.3 SD 14.5	63.5 SE 2.6 SD 16.0	62.9 SE 2.6 SD 16.0	62.8 SE 2.9 SD 17.9	65.0 SE 2.4 SD 14.9
Other L1s (n=36) 49.3%	60.5 SE 1.9 SD 11.7	62.3 SE 2.4 SD 14.6	59.0 SE 1.9 SD 11.5	57.9 SE 2.3 SD 14.2	66.6 SE 2.2 SD 13.4
t-test Results					
Levene's test	.274	.518	.161	.189	.404
Sig					
Sig (two-tailed)	.552	.733	.233	.198	.647
t	.598	.343	1.203	1.298	-.460
df	71	71	71	71	71
Effect Size (Cohen's *d* and Pearson's *r*) NOT APPLICABLE					

About half of Ghanaian test takers (50.7%) state English as their L1. According to descriptive statistics there is a small difference in the mean PTEA scores (in all score categories) of L1 English test takers compared to Other L1 test takers (English L1 test takers having 1-5 GSE points more than L1 Other test takers across the score categories). Other L1 test takers perform slightly better in their mean writing scores. Looking at CEFR levels, Other L1 test takers display the same level of performance as English L1 test takers (at B2) in all score categories except speaking where they are at B1+ level. The t-test suggests there are no statistically significant differences between the two groups in any score category. This suggests that Ghanaian test takers perform quite well on the test (B2 level) regardless of whether they have English as an L1.

8. Nigeria (ANOVA)

Table 38: Nigeria ANOVA for English L1 vs. Yoruba/Igbo/Hausa L1

PTEA Score Categories & Scores (Nigeria n=500)					
Group+	Overall	Listening	Reading	Speaking	Writing
English L1 (n=351) 70.2%	56.7 SE 0.7 SD 13.9	59.1 SE 0.8 SD 15.3	57.2 SE 0.7 SD 14.3	51.4 SE 0.9 SD 17.2	62.7 SE 0.7 SD 14.1
Yoruba (n=58) 11.6%	48.2 SE 1.9 SD 14.6	48.2 SE 2.1 SD 16.4	50.3 SE 2.1 SD 16.4	43.3 SE 2.1 SD 16.1	52.6 SE 1.9 SD 14.5
Igbo (n=36) 7.2%	57.2 SE 2.4 SD 14.6	59.5 SE 2.7 SD 16.2	59.3 SE 2.4 SD 14.9	47.8 SE 2.5 SD 15.5	64.4 SE 2.6 SD 15.7
Hausa ++ (n=26) 5.2%	57.2 SE 2.5 SD 13.1	60.2 SE 2.8 SD 14.4	58.1 SE 2.4 SD 12.4	51.6 SE 3.2 SD 16.4	63.0 SE 2.6 SD 13.2
ANOVA					
Levene's test Sig	.811	.546	.450	.530	.703
Sig (p)	0.00	0.00	.005	.007	.000
Post Hoc (Tukey)					
Yoruba mean difference (sig)	English –8.5 (0.00) Hausa –9.0 (.033) Igbo –9.0 (.013)	English –10.8 (0.00) Hausa –12.0 (.006) Igbo –11.3 (.003)	English –6.9 (.005) Igbo –9.0 (.020)	English –8.0 (.005)	English -10.0 (0.00) Hausa –10.3 (.012) Igbo –11.7 (.001)
Effect Size (eta squared)					
Eta Sq.	0.039 (S)	0.052 (SM)	0.026 (S)	0.025 (S)	0.054 (SM)

+ 5.8% (n=29) of the entire sample belonged to language groups other than the four above and are not included here.

++ Hausa were included even though strictly-speaking there was not enough (<30) to be included on the ANOVA

A large majority of Nigerian test takers (70.2%) state English as their L1. However, according to descriptive statistics, there is a slight difference in the mean PTEA scores (in all score categories) of L1 English test takers compared to other L1 test takers. All L1 groups (bar Yoruba) score between 50-60 on the GSE in all score categories except writing. All L1

groups score highest in their writing mean scores. Igbo and Hausa L1 test takers score slightly higher than L1 English test takers in all score categories apart from speaking where Igbo L1s score lower than English L1s. Yoruba L1s have a lower mean score than all other groups in all score categories. Comparing the CEFR levels, Other L1 test takers (Igbo and Hausa) display the same level of proficiency as English L1 test takers (at B2) in listening and writing. Igbo L1s have the highest mean score in reading, at just above 59. However, the ANOVA and Tukey tests reveal that the only statistically significant difference in scores ($p<.05$ level) is between Yoruba test takers and all others in overall, listening and writing; Yoruba and English/Igbo in reading, and Yoruba and English only in speaking. The effect sizes are small (eta sq = .01) for overall, reading and speaking, and close to medium (eta sq = .06) for listening and writing. The descriptive and inferential results suggest that L1 is not a significant factor in test performance in Nigeria unless the test taker is a Yoruba L1 speaker in which case they perform significantly lower than other L1 groups.

9. Cameroon

This data set for Cameroon includes only descriptive statistics as there were not enough participants in each L1 group with which to run a reliable t-test (>30).

Table 39: Cameroon descriptive stats for English L1 vs. French L1

PTEA Score Categories & Scores (Cameroon n=52)					
Group	Overall	Listening	Reading	Speaking	Writing
English L1	55.9	57.8	55.0	52.8	60.8
(n=28)	SE 2.1	SE 2.4	SE 2.5	SE 2.5	SE 2.5
53.8%	SD 11.1	SD 12.8	SD 13.2	SD 13.5	SD 13.5
French L1	49.7	50.5	51.9	41.2	56.9
(n=24)	SE 2.3	SE 2.4	SE 2.5	SE 2.1	SE 2.6
46.2%	SD 11.5	SD 11.8	SD 12.5	SD 10.5	SD 12.7

The descriptive statistics suggest that there is a difference between the test performance of English L1 test takers compared to that of French L1 test takers from Cameroon, suggesting that English L1 test takers perform better on the PTEA than their L1 French-speaking compatriots, with the biggest gap of 11 points on the GSE in the speaking score category. However, without a larger sample this difference in mean scores cannot be confirmed as statistically significant.

10. South Africa

Table 40: South Africa t-test for English L1 vs. Afrikaans L1

PTEA Score Categories & Scores (South Africa n=393)					
Group+	Overall	Listening	Reading	Speaking	Writing
English L1 (n=317) 80.7%	76.5 SE 0.7 SD 13.6	77.3 SE 0.7 SD 13.6	78.6 SE 0.7 SD 13.6	75.9 SE 1.0 SD 18.2	77.7 SE 0.7 SD 13.1
Afrikaans L1 + (n=69) 17.5%	74.8 SE 1.5 SD 13.2	74.9 SE 1.6 SD 13.5	77.2 SE 1.5 SD 13.1	72.8 SE 2.1 SD 17.5	76.1 SE 1.5 SD 12.7
t-test Results					
Levene's test	.717	.738	.620	.851	.899
Sig					
Sig (two-tailed)	.331	.179	.416	.198	.351
t	.973	1.347	.815	1.288	.934
df	384	384	384	384	384
Effect Size (Cohen's *d* and Pearson's *r*) NOT APPLICABLE					

+ 1.8% (n=7) of the entire sample belonged to language groups other than English and Afrikaans and are not included here.

A large majority of South African test takers (80.7%) state English as their L1. According to descriptive statistics, there is only a slight difference in the mean PTEA scores of L1 English test takers compared to Afrikaans L1 test takers in all score categories, but only by 1-3 GSE points in each category. English L1 test takers score slightly higher than L1 Afrikaans test takers in all score categories. In terms of CEFR levels, L1 Afrikaans test takers display a similarly high level of proficiency as English L1 test takers (at C1+) in reading and writing,

and B2 in speaking. The t-test reveals no statistically significant difference in the mean PTEA scores in any score category between South Africans with English as an L1 and Afrikaans as an L1.

11. Southern Africa

Table 41: Southern African t-test for English L1 vs. Other L1

PTEA Score Categories & Scores (Southern Africa n=156)					
Language	Overall	Listening	Reading	Speaking	Writing
English L1 (n=121) 77.6%	64.3 SE 1.5 SD 16.8	66.3 SE 1.5 SD 17.4	64.9 SE 1.5 SD 16.8	61.6 SE 2.1 SD 23.1	68.9 SE 1.4 SD 15.7
Other L1 + (n=35) 22.4%	58.3 SE 2.2 SD 13.4	59.3 SE 2.3 SD 13.8	60.4 SE 2.2 SD 13.5	51.8 SE 3.1 SD 18.5	64.9 SE 2.0 SD 12.1
t-test Results					
Levene's test	.045*	.093	.140	.018*	.136
Sig					
Sig (two-tailed)	.031	.030	.146	.011	.167
t	2.209	2.190	1.461	2.600	1.390
df	68.005	154	154	67.903	154
Effect Size (Cohen's *d* and Pearson's *r*)					
Cohen's *d*	0.535 (M)	0.352 (S)	n/a	0.631 (M)	n/a
r	0.258 (SM)	0.173 (S)	n/a	0.300 (M)	n/a

***Equal Variances not assumed**

A large majority of Southern African test takers (77.6%) state English as their L1. According to descriptive statistics, English L1 test takers score higher than L1 Other test takers in all score categories with a 4-6 point difference on the GSE in all score categories and a 10-point difference in speaking. Looking at CEFR levels, L1 Other test takers are at the same level as English L1 test takers (B2) in all categories apart from speaking (B1+) (in the overall category L1 Other is very close to B2). The t-test reveals no statistically significant difference in the mean PTEA scores between the groups in reading and writing. However, there is a

statistically significant difference in the scores between the groups in overall, listening and speaking scores. The magnitude of the significance is medium in overall and in speaking but small in listening (Cohen's *d* and Pearson's *r*).

12. Bangladesh

Table 42: Bangladesh t-test for English L1 vs. Other L1

PTEA Score Categories & Scores (Bangladesh n=500)					
Group+	Overall	Listening	Reading	Speaking	Writing
English L1 (n=111) 22.2%	46.3 SE 1.0 SD 11.4	47.1 SE 1.1 SD 12.6	45.4 SE 1.0 SD 11.3	47.7 SE 1.4 SD 15.4	48.2 SE 1.1 SD 11.6
Bengali L1 (n=384) 76.8%	45.4 SE 0.4 SD 9.7	45.6 SE 0.4 SD 9.6	45.3 SE 0.5 SD 10.6	46.1 SE 0.6 SD 13.4	47.5 SE 0.4 SD 9.4
t-test Results					
Levene's test Sig	.582	.035*	.809	.381	.051
Sig (two-tailed)	.397	.265	.938	.286	.512
t	.848	1.119	.078	1.069	.656
df	493	149.151	493	493	493
Effect Size (Cohen's *d* and Pearson's *r*) NOT APPLICABLE					

***Equal Variances not assumed**

+ a very small number of the sample (n=5) were neither English nor Bengali L1 and not used in this test.

Just under a quarter of Bangladeshi test takers (22.2%) state English as their L1. According to descriptive statistics, there is only a very slight difference in the mean PTEA scores of L1 English test takers compared to Bengali L1 test takers in all score categories (no more than a 2-point GSE difference) with English L1 test takers scoring only slightly higher in all score categories. Looking at the CEFR levels, Bengali L1 test takers are at the same lower level as English L1 test takers (B1) in all score categories. The t-tests reveal that there are no statistically significant differences in mean PTEA scores between L1 English test takers and L1 Bengali test takers in any of the score categories.

13. India (ANOVA)

Table 43: India ANOVA for English L1 vs. Telugu/Hindi/Punjabi/Tamil L1

PTEA Score Categories & Scores (India n=500)					
Group+	Overall	Listening	Reading	Speaking	Writing
English L1 (n=63) 12.6%	61.8 SE 2.2 SD 17.6	62.7 SE 2.2 SD 18.1	61.1 SE 2.3 SD 18.2	62.2 SE 2.6 SD 21.3	63.2 SE 2.1 SD 16.9
Telugu L1 (n=126) 25.2%	53.9 SE 1.0 SD 11.2	54.3 SE 1.0 SD 11.6	52.6 SE 1.0 SD 11.6	58.6 SE 1.4 SD 16.7	54.3 SE 0.9 SD 10.4
Hindi L1 (n=111) 22.2%	58.4 SE 1.3 SD 14.1	58.4 SE 1.3 SD 14.5	58.3 SE 1.3 SD 14.6	62.0 SE 1.8 SD 19.2	58.6 SE 1.2 SD 13.0
Punjabi L1 (n=77) 15.4%	48.6 SE 1.4 SD 12.3	49.1 SE 1.5 SD 13.2	48.6 SE 1.2 SD 11.3	47.7 SE 1.8 SD 16.2	51.7 SE 1.2 SD 11.2
Tamil L1 (n=34) 6.8%	58.9 SE 2.2 SD 12.9	59.2 SE 2.4 SD 14.0	58.7 SE 2.1 SD 12.6	60.4 SE 3.1 SD 18.4	60.3 SE 2.1 SD 12.4
ANOVA					
Levene's Test Sig	.001*	.002*	.000*	.005*	0.00*
Sig. (Welch)	.000	.000	.000	0.00	0.00
Sig. (Brown-Forsythe	.000	.000	.000	0.00	0.00
Post Hoc (Tukey)					
Punjabi mean difference (sig p=0.05)	English –13.291 (.000) Hindi –9.862 (.000) Tamil –10.314 (.002)	English -13.593 (.000) Hindi –9.282 (.000) Tamil –10.037 (.005)	English –12.498 (.000) Hindi -9.727 (.000) Tamil -10.197 (.003)	English –14.569 (.000) Hindi –14.299 (.000) Tamil –12.710 (.007) Telugu -10.957 (.000)	English -11.491 (.000) Hindi -6.869 (.003) Tamil -8.603 (.009)
Telugu mean difference (sig p=0.05)	English –7.968 (.001)	English -8.429 (001)	English –8.460 (.001) Hindi –5.689 (.013)	Punjabi 10.957 (.000)	English -8.921 (.000)

	Effect Size (eta squared)				
Eta Sq.	0.095 (M)	0.088 (M)	0.093 (M)	0.077 (M)	0.086 (M)

***Equal Variances not assumed**

+17.8% (n=89) of the entire sample belonged to many language groups other than these five and are not used in this test.

A small proportion of Indian test takers in the sample (12.6%) stated English as their L1. According to descriptive statistics, there is a difference in the mean PTEA scores of L1 English test takers compared to other L1 test takers, with English L1 test takers scoring higher in all score categories in the low 60s on the GSE. Tamil and Hindi L1 test takers are not far behind with a mean score above 58 in all score categories, with Telugu L1 test takers slightly lower at between 52 and 58 in all categories. Punjabi speakers are noticeably lower with a mean score of no more than 51 in any score category.

Looking at CEFR levels, English L1 test takers are at B2 in all score categories. L1 Hindi and Tamil test takers are at or very nearly at 59 - the same B2 level as English L1 test takers in all score categories. Telugu test takers are just behind at B1+ in all score categories and Punjabi test takers are lower at B1 in all score categories except writing (B1+). The ANOVA and Post Hoc test (Tukey) reveals that differences in mean scores between English, Hindi and Tamil L1 test takers are not statistically significant but that those between Punjabi and English (and Punjabi and other languages except Telugu) are statistically significant across the score categories. Telugu L1 test takers also demonstrate statistically significant differences in mean scores compared to English L1 test takers except in speaking. The magnitude of the effect is medium in all cases.

14. Malaysia (ANOVA)

Table 44: Malaysia ANOVA for English L1 vs. Malay & Chinese L1

PTEA Score Categories & Scores (Malaysia n=500)					
Group+	Overall	Listening	Reading	Speaking	Writing
English L1 (n=152) 30.4%	69.7 SE 1.1 SD 14.2	71.1 SE 1.2 SD 14.9	69.9 SE 1.1 SD 14.3	67.9 SE 1.5 SD 18.5	72.2 SE 1.0 SD 13.0
Malay L1 (n=157) 31.4%	56.8 SE 0.9 SD 11.8	57.9 SE 1.0 SD 13.1	58.3 SE 0.9 SD 11.7	50.4 SE 1.2 SD 15.0	61.2 SE 0.9 SD 12.3
Chinese Languages L1 (n=181) 36.2%	54.1 SE 0.9 SD 12.5	54.5 SE 1.0 SD 13.8	55.4 SE 1.0 SD 13.6	48.8 SE 1.0 SD 13.9	58.1 SE 1.0 SD 13.6
	ANOVA				
Levene's test Sig	.023*	.054	.015*	.000*	.173
Sig (p)	.000 (Welch & Brown Forsythe)	.000	.000 (Welch & Brown Forsythe)	.000 (Welch & Brown Forsythe)	.000
	Post Hoc (Tukey)				
English mean difference (sig p=0.05)	Chinese 15.611 (.000) Malay 12.877 (.000)	Chinese 16.609 (.000) Malay 13.207 (.000)	Chinese 14.526 (.000) Malay 11.546 (.000)	Chinese 19.071 (.000) Malay 17.476 (.000)	Chinese 14.056 (.000) Malay 10.968 (.000)
	Effect Size (eta squared)				
Eta Sq.	0.214 (S)	0.206 (S)	0.179 (S)	0.224 (S)	0.173 (S)

***Equal Variances not assumed**

+ 2% (n=10) of the Malaysian sample belonged to language groups other than these three and not used in this test.

About a third of Malaysian test takers (30.4%) state English as their L1 and the other two-thirds are split fairly equally between Malay and Chinese languages. According to descriptive statistics, there is a difference in the mean PTEA scores of L1 English test takers compared to Malay and Chinese L1 test takers, with English L1 test takers scoring a lot higher on the GSE in all score categories (high 60s to early 70s). Malay and Chinese L1 test takers are below 59 in almost every score category in comparison. Looking at CEFR levels, English L1 test takers

are at B2 in all score categories whilst Chinese and Malay are at B1+ in overall, listening and reading and B1 in speaking. L1 Malay test takers match L1 English test takers in writing (B2) but are still at the lower end of the B2 band compared to L1 English test takers. The ANOVA and Tukey test revealed that there are statistically significant differences between English L1 test takers and L1 Malay and Chinese in all score categories; however, the effect size for the statistics is small in all score categories.

15. Pakistan

Table 45: Pakistan t-test for English L1 vs. Urdu L1

PTEA Score Categories & Scores (Pakistan n=500)					
Group+	Overall	Listening	Reading	Speaking	Writing
English L1	50.5	50.5	50.5	55.2	50.4
(n=80)	SE 1.4	SE 1.4	SE 1.5	SE 2.0	SE 1.3
16%	SD 13.3	SD 13.0	SD 14.2	SD 18.5	SD 12.1
Urdu L1	48.4	48.9	47.8	50.7	49.9
(n=369)	SE 0.6	SE 0.6	SE 0.6	SE 0.8	SE 0.6
73.8%	SD 12.1	SD 12.6	SD 12.5	SD 16.6	SD 11.5
	t-test Results				
Levene's test	.245	.374	.121	.167	.364
Sig					
Sig (two-tailed)	.186	.297	.087	.035	.775
t	1.325	1.043	1.716	2.119	.287
df	447	447	447	447	447
	Effect Size (Cohen's *d* and Pearson's *r*)				
Cohen's *d*	n/a	n/a	n/a	0.200 (S)	n/a
r	n/a	n/a	n/a	0.099 (M)	n/a

+ 10.2% (n=51) of the Pakistani sample belonged to multiple language groups other than these two and were not used in this test.

The descriptive statistics suggests that Pakistani test takers have a low proficiency in English no matter what their L1 – either English or Urdu (other L1 groups of Pashto and Punjabi L1 speakers also scored similarly across the score categories but were less than 30 in sample size). The GSE scores tend to be at or around 50 on the GSE for both groups in all score

categories. The CEFR level is at B1 in every score category. The t-tests indicate that there is a statistically significant difference in scores between L1 English and L1 Urdu test takers in only the speaking score category, with a small effect size in Cohen's *d* but medium in Pearson's *r*.

16. Philippines

Table 46: Philippines t-test for English L1 vs. Tagalog L1

PTEA Score Categories & Scores (Philippines n=276)					
Group+	Overall	Listening	Reading	Speaking	Writing
English L1	59.9	60.9	60.6	57.7	64.2
(n=90)	SE 1.8	SE 1.7	SE 1.7	SE 2.2	SE 1.6
32.6%	SD 17.1	SD 17.0	SD 16.9	SD 21.2	SD 15.9
Tagalog L1	58.0	58.8	58.9	57.4	61.0
(n=161)	SE 1.1	SE 1.2	SE 1.1	SE 1.4	SE 1.1
58.3%	SD 14.8	SD 15.9	SD 14.5	SD 18.1	SD 15.2
t-test Results					
Levene's test	.045*	.229	.076	.023*	.300
Sig					
Sig (two-tailed)	.382	.326	.387	.922	.112
t	.876	.984	.867	.098	1.595
df	163.210	249	249	160.985	249
Effect Size (Cohen's *d* and Pearson's *r*) NOT APPLICABLE					

+ 9.1% (n=25) of the Filipino sample belonged to multiple language groups other than these two and were not used in this test.

The descriptive statistics suggest that there is a very slight difference between English L1 test takers and Tagalog L1 test takers in their PTEA test scores, with both groups having mean scores between a very small range of 57 and 61 on the GSE in all score categories except writing. The t-test confirms that there is no statistical significance in these differences in any of the score categories. Filipino test takers perform similarly well on the PTEA, whether they are Tagalog L1 or English L1 speakers, and generally at or just below the CEFR B2 level.

17. Sri Lanka (ANOVA)

Table 47: Sri Lanka ANOVA for English L1 vs. Sinhalese & Tamil L1

PTEA Score Categories & Scores (Sri Lanka n=197)					
Group+	Overall	Listening	Reading	Speaking	Writing
English L1 (n=47) 23.8%	52.9 SE 2.3 SD 16.2	53.7 SE 2.4 SD 16.6	51.5 SE 2.4 SD 16.8	54.7 SE 2.7 SD 19.1	54.4 SE 2.2 SD 15.6
Sinhalese L1 (n=77) 39%	47.5 SE 1.2 SD 10.5	47.3 SE 1.2 SD 11.0	48.8 SE 1.2 SD 10.8	47.6 SE 1.7 SD 15.2	48.7 SE 1.1 SD 10.4
Tamil L1 (n=68) 34.5%	46.4 SE 1.1 SD 9.6	46.7 SE 1.1 SD 9.3	46.2 SE 1.2 SD 10.2	47.7 SE 1.9 SD 16.1	48.1 SE 1.1 SD 9.3
ANOVA					
Levene's test Sig	.000*	.000*	.000*	.068	.000*
Sig (p)				.044	
Welch (sig)	.058	.035	.110	n/a (.080)	.046
Brown-Forsyth (sig)	.024[8]	.011	.113	n/a (0.53)	.016
Post Hoc (Tukey)					
English L1 mean difference (sig)	Sinhalese 5.382 (.041) Tamil 6.430 (.014)	Sinhalese 6.441 (.013) Tamil 7.001 (.008)		Although not violating the assumption of equal variances assumed (.068 in Levene's test) and significant in the ANOVA (at.044) the Welch test is more accurate with an indication of 'no significance' (.080)	Sinhalese 5.788 (.021) Tamil 6.357 (.012)
Effect Size (eta squared)					
Eta Sq.	0.044 (S)	0.054 (S)	n/a	n/a	0.049 (S)

*** Equal Variances not assumed**

+2.5% (n=5) of the Sri Lankan sample belonged to language groups other than these three and not used in this test.

[8] Here, only the Brown-Forsyth test indicates significance.

The descriptive statistics shows that Sri Lankan test takers with English as an L1 have higher mean scores in all score categories (between 51 and 55 on the GSE) than either Sinhalese or Tamil L1 test takers (both between 46 and 49 in all score categories). The difference in CEFR bands is that of B1+ for English L1 test takers compared to B1 for Tamil and Sinhalese L1 test takers in all score categories. The ANOVA and Tukey results indicate that there is a statistically significant difference in scores between L1 English and L1 Tamil and Sinhalese L1 test takers in the overall, listening and writing score categories only, with a small effect size in each case.

Summary of Findings

Most of the tests conducted on the larger data sets indicate that L1 English test takers display higher mean scores than Other L1 test takers, with statistically significant differences. A large effect size is evident when comparing the OC and EC in this way. In terms of smaller regions and country groups, the results indicate that OC African test takers in particular display a high proficiency reflected in their PTEA test scores and that this is not necessarily dependent on their L1 being English. Nigeria is notable for slightly lower mean test scores for its English L1 test takers in comparison to test takers with other Nigerian L1s.

Apart from Nigeria and Cameroon, even where mean test scores for English L1 speakers are higher than Other L1s and differences in mean scores are statistically significant in some score categories (East and Southern Africa), outer circle African test takers tend to score at a high level on the PTEA in all or most categories (B2) no matter what their L1 is (English or other). In OC Asia the Philippines also seems to follow this pattern as does India to some extent with Hindi and Tamil. For some OC Asian countries such as Bangladesh and Pakistan,

having a stated L1 of English does not seem to account for high mean test scores at all and the mean scores are similar to those with other L1s.

4.1.3 Test Data from Other Academic English Tests

4.1.3.1 Data Collection

Further evidence of outer circle test taker performance on academic English tests as proof of proficiency was gained by viewing the test performance data from other academic English tests. Strictly-speaking this is documentary data as I did not have access to the original statistics and one important missing piece of data was the size of sample the data sets were based on. Cambridge Assessment (CA) and ETS does not display this information on their websites.The Cambridge Assessment (CA) website (UCLES, 2015) displays publicly-available data on test score averages categorised by country. Among all the potential tests such as First Certificate in English (FCE), Cambridge Advanced Certificate (CAE) and Cambridge Proficiency in English (CPE), only IELTS was available and suitable for analysis. On the IELTS website (IELTS Partners, 2013) there is a summary, for each year, of the mean test scores for various countries. The same type of data was retrieved for TOEFL scores for the year 2013 (Educational Testing Service, 2014) and the information on the ETS website concording the TOEFL scores with the CEFR levels was consulted (Educational Testing Service, 2015).

4.1.3.2 Analytic Approach

IELTS and TOEFL-iBT scores were converted into the CEFR band scores and compared with the PTEA mean test scores as per country. Tables showing the highest ranked countries in terms of mean test scores for overall score and in each of the four skills was displayed in

tables. From the data it can be seen how outer circle countries compare against expanding circle countries in mean test scores according to CEFR level. This data was useful as a comparison with the PTEA data because an assumption is that they should indicate similar proficiencies for the same countries. If they don't, this could call into question the validity of one of the tests although there could be other explanations such as the normality of the samples or an indication that certain academic tests are taken by students with differing proficiencies from the same country. Evidence from other tests supports or contradicts my findings from the quantitative PTEA data and hence serves as one kind of triangulation.

4.1.3.3 Findings - IELTS Mean Score Data (2013)

To act as a comparison to the PTEA 2014 test taker data, I consulted the IELTS website for data of a similar nature on nationality and test taker score profiles "These figures show the mean overall and individual band scores achieved by 2013 Academic and General Training candidates from the top 40 places of origin" (IELTS Partners, 2013). I accessed the data for 2013 and in terms of the typicality of these figures, IELTS state "These figures are broadly in line with statistics for previous years" (ibid). However, a mean average is given with no reference to the sample size (*n*) of each nationality. The data from the wesbite is not as comprehensive as my PTEA data in terms of covering the same number of countries. In terms of the CEFR levels, for comparability with the PTEA data IELTS align their test to the CEFR as shown in table 48 (IELTS, 2016).

Table 48: Key for Colours and CEFR Levels (IELTS)

Colour Key	CEFR Level
Blue	C1+
Green	B2
Yellow	B1

Data on mean scores was found for only seven outer circle countries as follows:

Table 49 Average IELTS Scores for Outer Circle (2013)

Country	Overall	Listening	Reading	Writing	Speaking
Malaysia	6.8	7.2	7.0	6.1	6.6
Philippines	6.8	7.1	6.7	6.2	6.9
Nigeria	6.4	6.3	6.0	6.2	6.8
Sri Lanka	6.4	6.6	6.2	6.0	6.7
Hong Kong	6.4	6.8	6.7	5.8	6.1
India	6.0	6.2	5.8	5.6	5.9
Bangladesh	5.9	5.9	5.7	5.7	6.0

Analysing the IELTS data in terms of CEFR bands, it can be seen that the outer circle countries display no mean scores below B2 in any score category (table 49) with Malaysia and the Philippines having the highest overall means at 6.8 each. There is only data for one OC African country (Nigeria) which displays a higher mean than in the PTEA at B2 in every score category and C1+ for speaking. Data for mean scores was found for a selection of expanding circle countries listed below ranked from highest to lowest:

Table 50 Average IELTS Scores for Expanding Circle (2013)

Country	Overall	Listening	Reading	Writing	Speaking
France	6.7	6.9	7.1	6.0	6.6
Russia	6.5	6.7	6.6	5.9	6.6
Italy	6.5	6.5	7.0	5.9	6.4
Mexico	6.4	6.5	6.7	5.8	6.4
Brazil	6.3	6.2	6.7	5.9	6.3
South Korea	5.9	6.1	6.1	5.5	5.7
China	5.7	5.8	6.0	5.3	5.4
Japan	5.7	5.8	6.0	5.3	5.5
Turkey	5.5	5.4	5.6	5.1	5.5
Saudi Arabia	5.0	4.9	4.8	4.6	5.6

As with the PTEA mean scores per country (table 23), the same middle eastern and east Asian examples display the lower mean scores (South Korea, China, Japan, Turkey and Saudi Arabia) but the level indicated by the CEFR is higher than on the PTEA at B2 and B1+. In comparing indications of proficiency, the IELTS data contradicts my quantitative data in that many countries in the expanding circle (Brazil, Turkey, China and Japan) are all at B2 overall according to IELTS scores but in the PTEA 2014 sample these are all at B1 or A2 level. In

terms of the CEFR, the IELTS data suggests that test takers from various countries have a higher proficiency in English than that indicated by the PTEA.

4.1.3.4 Findings - TOEFL-iBT Mean Score Data (2013)

Educational Testing Service produces the TOEFL-iBT (internet-based test) and displays the alignment of the TOEFL scores with the CEFR on its website (see table 51 below). It also produces comprehensive data for mean test scores as per nationality of the test taker. The data below summarises which countries were top in terms of mean scores in each continent for the year 2013 (Educational Testing Service, 2014). Unfortunately, as with IELTS, ETS does not provide information as to the sample number from which the data was drawn but countries with less than 30 test takers are excluded from the data (usually small island nations in the Pacific and Indian Oceans and the Caribbean). The alignment of TOEFL-iBT scores with the CEFR is expressed in table 51 below and is taken from the ETS website (Educational Testing Service, 2015).

Table 51 TOEFL-iBT Scores and Alignment with the CEFR

CEFR level	Reading (0–30)	Listening (0–30)	Speaking (0–30)	Writing (0–30)	Total (0–120)
C1 or above	24	22	25	24	95
B2	18	17	20	17	72
B1	4	9	16	13	42
A2	n/a	n/a	10	7	n/a

The data for each continent is displayed and summarised below using the following colour key:

Table 52: Colour Key for TOEFL-iBT Mean Scores Tables

Colour Key	CEFR Level
Blue	C1+
Green	B2
Yellow	B1

In addition to the above key, a highlighted country name in grey indicates an outer circle country.

Africa

Table 53 Average TOEFL-iBT Scores for Africa

Country	Reading	Listening	Speaking	Writing	Overall
1. South Africa	22	25	26	24	98
2. Mauritius	23	24	23	24	93
3. Zimbabwe	21	22	24	24	91
4. Botswana	20	21	23	23	88
5. Namibia	18	22	24	22	87
6. Swaziland	18	21	23	22	84
6. Zambia	18	20	24	22	84
8. Kenya	19	20	22	22	83
8. Egypt	19	21	22	21	83
10. Malawi	18	19	22	22	80
10. Madagascar	19	20	21	21	80
10. Tunisia	19	20	21	20	80
10. Uganda	18	19	22	21	80
14. Morocco	18	20	21	20	79
15. Ghana	18	19	21	21	78
15. Nigeria	18	19	21	21	78
15. Reunion	19	19	20	20	78
18. Eritrea	18	19	20	20	77
18. Sierra Leone	17	19	22	20	77
20. Ethiopia	17	19	21	19	76
21. Sudan	16	19	21	19	75
22. Tanzania	16	17	21	19	74
22. Somalia	16	18	21	19	74
24. Algeria	16	18	20	18	73
24. Libya	16	18	21	18	73
26. Rwanda	15	17	20	19	72
27. Cameroon	16	16	20	19	70
27. Mozambique	15	17	20	19	70
29. Burundi	15	16	20	18	69
29. Chad	15	16	20	18	69
29. Liberia	14	16	21	18	69
29. Gabon	15	16	19	18	69
29. Niger	15	17	19	18	69
19 more non-Anglophone African countries					Below B2

Source: (Educational Testing Service, 2014)

Outer circle African countries are clustered at or near the top of the table (B2+). This indicates that within Africa, Anglophone or outer circle countries display higher TOEFL-iBT mean scores than expanding circle African countries. Confirming my PTEA data, South Africa ranks highest within outer circle Africa in terms of mean scores and Nigeria and Cameroon score lower down the table than other Anglophone African countries. Other outer circle African countries not displayed were: Gambia (67 overall score); Seychelles (no data); Lesotho (no data). All other 19 Francophone, Lusophone and other African nations not listed above scored a mean average overall score of below B2.

Asia

Table 54 Average TOEFL-iBT Scores for Asia

Country	Reading	Listening	Speaking	Writing	Total
1. Singapore	24	25	24	25	98
2. Israel	22	25	24	22	93
3. India	22	23	23	23	91
4. Pakistan	21	22	24	23	90
5. Malaysia	22	23	22	23	89
5. Philippines	21	22	24	23	89
5. Lebanon	20	22	23	22	89
8. Sri Lanka	20	21	22	21	85
8. S. Korea	22	21	21	22	85
10. Bangladesh	20	21	21	22	84
10. Bahrain	18	21	23	21	84
12. Hong Kong	19	21	21	22	83
12. Nepal	20	20	21	21	83
14. Indonesia	20	21	20	21	82
14. N. Korea	20	20	20	21	82
14. Iran	20	20	21	21	82
17. Kazakhstan	18	20	22	20	80
18. Myanmar	18	19	20	21	79
18. Taiwan	20	20	20	20	79
18. Bhutan	17	18	22	21	79
18. Uzbekistan	18	19	21	20	79
22. Azerbaijan	18	19	20	20	78
22. Vietnam	19	19	19	21	78
22. Jordan	18	20	21	19	78
25. China	20	18	19	20	77
25. Qatar	17	20	21	19	77
27. Syria	17	19	21	19	76
27. UAE	17	19	21	19	76

27. Kyrgystan	17	19	21	20	76
27. Macao	18	18	19	20	76
27. Thailand	18	19	19	20	76
32. Palestine	17	19	21	19	75
33. Turkmenistan	16	18	20	19	74
34. Kuwait	16	18	20	18	72
35. Afghanistan	15	16	21	19	71

Source: (Educational Testing Service, 2014)

Again, as with the African countries, the outer circle Asian countries are clustered at or near the top of their continental table (B2+) with Singapore the highest ranking as with the 2014 PTEA data. This indicates that within Asia, Anglophone or outer circle countries display higher TOEFL-iBT mean average scores than expanding circle Asian countries. However, a large number of expanding circle countries display a high proficiency in English (B2 overall) as indicated by their overall TOEFL-iBT scores (from Indonesia to Kuwait) and countries such as Pakistan, Sri Lanka, and Hong Kong who ranked low in terms of mean PTEA scores actually have high mean scores on the TOEFL-iBT (relating to B2 in the CEFR).

Europe

Table 55 Average TOEFL-iBT Scores for Europe

Country	Reading	Listening	Speaking	Writing	Overall
1. Austria	24	26	26	25	100
2. Denmark	23	25	26	24	98
3. Germany	23	25	25	24	97
3. Belgium	24	25	24	24	97
3. Luxembourg	23	24	25	24	97
3. Switzerland	24	25	24	24	97
7. Slovenia	23	25	24	24	96
7. Finland	23	25	24	24	96
9. Iceland	22	25	24	23	95
9. Portugal	23	25	24	23	95
11. Estonia	22	24	24	23	94
11. Norway	21	24	25	23	94
11. Sweden	21	25	25	23	94
28 more countries					All score B2 or above

Source: (Educational Testing Service, 2014)

According to the TOEFL-iBT data, Europe (all expanding circle apart from Malta and Cyprus) displays a much higher proficiency in English in the form of overall scores than outer circle countries in Africa and Asia, and a large number scored at C1+ overall. 28 further countries are not listed but they all scored B2 and above overall on the TOEFL-iBT. This matches some of the PTEA mean score data for the expanding circle (table 23) but the mean scores on the TOEFL-iBT suggest that the European expanding circle exceeds the proficiency of outer circle countries.

North & South America

Table 56 Average TOEFL-iBT Scores for North & South America

Country	Reading	Listening	Speaking	Writing	Total
1. Argentina	23	24	23	23	93
1. Costa Rica	23	24	23	23	93
1. Uruguay	23	24	23	23	93
4. Paraguay	21	22	22	22	87
4. Peru	21	22	22	22	87
4. Puerto Rico	20	22	23	22	87
7. Jamaica	19	21	24	22	86
7. Mexico	21	22	22	21	86
9. El Salvador	20	21	23	21	85
10. Nicaragua	19	21	22	21	84
10. Panama	19	21	22	21	84
11 more countries					B2 overall*

Source:(Educational Testing Service, 2014)

***11 further nations (not listed) scored B2 overall. Only 2 (Guadeloupe and Haiti) scored lower than B2 overall in the region.**

According to the TOEFL-iBT data, the Americas (expanding circle, apart from the Anglophone Caribbean) display a high proficiency in English in the form of overall scores and a large number scored C1+ overall in listening skills. Jamaica is considered an outer circle country even though it is exempt from SELTs in the UK and displayed a mean score of B2 overall.

Summary

Within their continents, outer circle African and Asian countries display a higher proficiency in English than their expanding circle counterparts. However, the TOEFL-iBT data for Europe and South America suggest that the proficiency of test takers from these expanding circle regions display equal if not higher levels of proficiency in English than the outer circle African and Asian countries and this serves as a triangulation with the PTEA mean score findings.

4.1.4 Perceptions of Proficiency and Ownership

4.1.4.1 The Outer Circle Student Survey

RQ3 (PERCEPTIONS OF PROFICIENCY-OWNERSHIP): What are stakeholder perceptions of outer circle proficiency and ownership of English for university entrance?

4.1.4.1.1 Data Collection

Data from the first case study tutorials and themes from the literature review were drawn on in the design of the outer circle student survey. This survey (appendix 15) took the form of an online questionnaire with some closed questions for biographical information and a variety of open-ended questions. These were designed to gather the attitudes and opinions of students from the outer circle on issues raised in the literature and first tutorials with the 4 cases regarding proficiency in English and ownership of English. Seven Likert scales[9] were also used within the survey to gauge opinions and these are included within the analysis of the relevant themes below. The findings from these scales were triangulated with the themes from analysis of the qualitative data in the same survey. The questions were mainly concerned with:

[9] To view the scales, go to www.warwick.ac.uk/rgwilson and click on the tab 'Outer Circle Student Survey'

- The participants' experiences of SELTs and CSELTs in university admissions.
- The participants' previous education in English.

From the open questions I obtained comments that could be coded in a thematic analysis for the purpose of compiling more qualitative evidence for comparison with the themes from the case study and recruitment survey in strand two (interpreted in the discussion chapter).

The survey was piloted with students from Warwick, and distributed using outer circle student society email addresses found on university student union websites, for example: 'X University Nigerian Society' and other contacts of the author to over 100 student society university email addresses in the UK. Paper A-4 posters were also displayed in universities in London and sent to various UK universities to encourage respondents.

4.1.4.1.2 Analytic Approach

After responses were gathered, data from the online survey was downloaded in Excel and put into Nvivo software for coding. Themes were developed as articulated in the methodology chapter above to produce a thematic description of the data from the online survey. The themes emerging from the responses are outlined below, illustrated with quotes and summarised at the end in a "thematic map". The identification number (ID) of the participant and their nationality is included after direct quotes in the thematic descriptions, for example "outer circle survey participant number 20 from Kenya" is shortened to (OC20 Kenya). Biographical and descriptive data is presented below in table 57.

4.1.4.1.3 Findings - Profiles of Participants

Eleven students replied to the emails sent out to the 100 or so country-categorised student societies at universities throughout the UK and through contacts at Warwick. Six individuals from Cameroon and Malawi then responded via an individual in a student society at the University of Edinburgh. The final sample for the outer circle survey in table 57 comprises the following seventeen individuals and reflects the background demographic data gathered for questions 1-6 in the survey.

Table 57: Profiles of Outer Circle Students in the survey (sorted alphabetically)

ID	Nationality	Gender	Year of Birth (Age)	L1	Other Languages spoken
OC1	Cameroon	Male	1970 (46)	Aghem	English, French, Other Cameroonian Languages
OC8	Hong Kong (China)	Male	1991 (25)	Cantonese	English, Mandarin
OC7	India	Female	1970 (46)	English	A little bit of French, Italian, Portuguese, Hindi, Marathi
OC17	Malawi	Female	1981 (35)	Chichewa	English
OC5	Nigeria	Female	1972 (44)	Igbo	English, Hausa
OC6	Nigeria	Female	1978 (38)	Tiv	English
OC4	Pakistan	Female	1978 (38)	Urdu	English, Punjabi, Hindi
OC9	Pakistan	Female	1975 (41)	Urdu	English
OC2	Singapore	Female	1994 (22)	English	French
OC16	Singapore	Female	1983 (33)	Chinese	English
OC3	South Africa	Male	1966 (50)	Setswana	English, Afrikaans, Sepedi, Sesotho, IsiXhosa, isiZulu
			Participants below were not studying in the UK+		
OC10+	Cameroon	Male	1954 (62)	Nweh	English, Pidgin English, French
OC12+	Cameroon	Male	1984 (32)	English	French
OC13+	Cameroon	Female	1982 (34)	Bali	French, English
OC15+	Cameroon	Male	1983 (33)	Bangwa	French, American Sign Language,Pidgin English, Bangwa-Lebialem
OC11+	Malawi	Male	1976 (40)	Chitumbuka	English, Chichewa and Sign Language
OC14+	Malawi	Male	1986 (30)	Chichewa	English

+These 6 participants were not studying in the UK but were working in their home countries so some of the questions on admissions and UK study did not apply to them. However, they were asked questions relating to their proficiency in English, their relationship to English and the value of their CSELTs.

The sample in table 57 includes an equal balance of genders (8 male and 9 female). All of the students surveyed were over the age of 21 and included a range of nationalities from Africa and Asia. A wide variety of first languages were spoken and many participants were multilingual. English was stated as an L1 by three of the 17 individuals.

Table 58 summarises questions 7 – 17 in the survey[10], which looked at the background English language qualifications of the respondents and information on their CSELTs regarding proof of proficiency at point of admission. The respondents are arranged alphabetically in terms of country of origin.

Table 58: Summary of OC Survey Questions 7-17, English Language Qualifications

Test Taker ID and Country	Qualification Sought in UK	Type of Proof of Proficiency used at point of Admission	Name of CSELT	Previous Qualifications	CSELT used instead of IAET for proof of proficiency
OC#1 (Cameroon)	MA Teaching English to Young learners	IELTS 8.0	Cameroon O Level (A)	BA in English MA and MPhil in English literary studies,	No
OC#8 (Hong Kong)	MSc Applied Linguistics and Second Language Acquisition	IELTS 8.5	Hong Kong HKALE (B)	--	No
OC#7 (India)	MA TESOL	IELTS 8.5/ 100/100 on an English language test conducted by the University of Leeds	My entire education in India, from school to MA, was in English	My entire education in India, from school to MA, was in English	Don't Know
OC#17 (Malawi)	PhD in Applied Linguistics	IELTS 7.5/sample of written work e.g. MA Thesis	Malawi MCSE (Grade 3)	--	No

[10] To view the instrument, go to www.warwick.ac.uk/rgwilson and click on the tab 'Outer Circle Student Survey'

OC#5 (Nigeria)	MA TESOL and ICT	My BEd (Language Arts:English) assessed at CEFR Level B2.	Nigerian WAEC/NECO (B)	B.Ed	Yes
OC#6 (Nigeria)	MSc Financial Management	BSc degree taught in English Language	Nigerian WAEC	BSc degree taught in English Language (2:2)	Yes
OC#4 (Pakistan)	BEd.	Email discussion with academic team and Pakistan University Confirmation letter of medium of education in English	social sciences taught in English (B)	BEd.	Yes
OC#9 (Pakistan)	PhD in ELT	IELTS 8.0/ a CSELT/a written assignment	Pakistan B.H.E (A)	--	No
OC#2 (Singapore)	BSc Biomedicine	GCSE (High School)	UK GCSE (A*) in Singapore	--	Yes
OC#16 (Singapore)	MA in ELT	UG degree from my country	Singapore O Level (Grade A)	--	Yes
OC#3 (South Africa)	MA in ELT	IELTS 8.0	South Africa National Senior Certificate (65%)	--	No
Participants below were were working in the OC at time of survey					
*OC#10 (Cameroon)	Assistant lecturer/school principal	+ GCE A Level English , University Compulsory English Language courses	Cameroon O Level (B)	PhD,	n/a
*OC#12 (Cameroon)	Technician	--	Cameroon O Level (C)	BA/BSc	n/a
*OC#13 (Cameroon)	Assistant Researcher	+ B and B+ in University compulsory courses Eng1 and Eng2	Cameroon O Level (B)	MSc Medical Microbiology and Parasitology	n/a
*OC#15 (Cameroon)	Teacher	+ GCE A/L English language and The University Use of English	Cameroon O Level (C)	BEd.	n/a

		courses 1 and 2			
*OC#11 (Malawi)	Inclusive Education District Coordinator	--	Malawi MSCE (6)	MA/MSc	n/a
*OC#14 (Malawi)	ICT Systems Engineer	--	Malawi MSCE (4)	BA/BSc	n/a

*** Instead of course of study, these participants put their current occupation**

+ They were also asked which existing English language proficiency qualifications apart from the CSELTs they had acquired.

Table 58 illustrates the admissions experiences of a sample of outer circle students entering UK universities. The participants can be seen to be well qualified and most were pursuing postgraduate studies (including two PhD students). The respondents had already obtained undergraduate qualifications either in their country of origin or in the UK. Five of the 11 studying in the UK said that their CSELT was used at point of admission instead of a SELT, five of the 11 said a SELT was used instead of their CSELT, and one out of 11 did not know which specific means were used to ascertain their English language proficiency for university. All had undergraduate degrees taught in English in their country of origin. The students faced an ambiguous situation in terms of admission policies in the light of their previous English qualifications and education in the medium of English. Six out of the seventeen students had taken an international academic English test (IAET) – IELTS in every case. Three of those scored 8.5 overall, two scored a band 8 and one obtained a band 7.5 overall indicating a high level of language proficiency.

4.1.4.1.4 Findings - Thematic Analysis of OC Survey Data

A thematic analysis of their online comments collected opinions and produced themes related to two broad areas: their proficiency in English (including ownership) and their admissions experiences into UK institutions. These are illustrated below.

1. Proficiency in English (including ownership)

This broad area comprised the following themes, illustrated one by one below:

- Medium of Education
- Historical Relationship with the UK
- Social Class
- English as an L1 in a Diverse Nation
- A Varied Range & Depth of Use of English
- Personal Impressions of Proficiency

"Medium of Education"

Many of the students indicated how they had been educated in the medium of English: *our language of teaching in Malawi school is English, which is standardized also* (OC14 Malawi); *it is the sole language of instruction from elementary education to Secondary through completion of High school and University studies* (OC15 Cameroon); *Mode of learning, teaching and speaking is matching those of the UK* (OC11 Malawi). However, in some outer circle territories not all students would have had an English medium education as mentioned by OC8 (Hong Kong).

"Historical Relationship with the UK"

The experience of English medium education often derives from systems inherited from colonial times: *The Cameroon General Certificate of Education board originates from the*

UK education system (OC10 Cameroon). Cameroon is an example of an outer circle nation split historically between different colonial powers (Britain, France and Germany) so that the English legacy is felt in a particular part of the country: *a significant part of my country is English speaking in the sense that its educational system is in the medium of English from the first day in school* (OC1 Cameroon). However there are some affronted feelings here to do with the shared colonial heritage in Cameroon: *I think the UK sees my country as a French colony and so ignores its own legacy in that country's educational system* (OC1 Cameroon).

"Social Class"

Social Class is seen as a factor in the inheritance of English for outer circle students in that a lot of students studying abroad will be from a middle-class background: *We were colonized by the British and most students who come here to further their studies are from comfortable families* (OC6 Nigeria). This phenomenon includes an urban dimension because in many cities a speaker's stated L1 may not be used as much as English: *Many students in cities are alienated from their mother tongues and English is their only language in daily use* (OC15 Cameroon). In Cameroon an elite will either be English or French-speaking or both: *the 'educated' citizens are either Bilingual or make claims of having at least one of these languages* (OC10 Cameroon).

"English as an L1 in a diverse nation"

The practicality of English as a means of communication in a diverse nation and adoption of it as a *de facto* L1 for many speakers was acknowledged: *for many children whose mothers may not come from the same L1 background like their fathers, English is the only language*

they speak so I don't see why English would not be their first language (OC1 Cameroon). This practical adoption of English is also seen in India: *for many people English is the only or the dominant language in their family, social and professional lives* (OC7 India) and also multi-ethnic Singapore: *the majority (96.1%) of Singaporeans speak English fluently and it's considered the primary language* (OC2 Singapore). In contrast, in Malawi, English does have an official role but in everyday life a local language still functions strongly as a lingua franca: *use of English is confined to official communication e.g. parliament, schools, top business meetings etc. In all social events and daily undertakings people use Chichewa*. (OC17 Malawi). In certain outer circle countries or territories such as Hong Kong (OC8) and Pakistan (OC9), participants from those two states did not view their nations as "English-speaking".

"A Varied Range & Depth of Use"

Connected to the above, among the responses there was a varied portrayal of the situations and depth of usage of English in a particular country. The example of one of the Malawian respondents (OC17 above) indicates an indigenous language (Chichewa) functioning as a social lingua franca with English remaining as a lingua franca in more formal official domains. However this view is contradicted by another survey respondent from the same country who was asked whether they thought Malawi was an English-speaking nation: *99%, because almost everything are communicated in English* (OC14 Malawi). English as a medium of communication in "official" domains is mirrored in the comments from the South African respondent: *English is a medium of business, government and education engagement* (OC3 South Africa). As mentioned above (OC1 Cameroon), in the domestic context in large multilingual African countries such as Cameroon, English is used at home as a "new" L1 in

families of mixed linguistic parentage but even in some other countries where "indigenous" L1s are still prominent in the home domain (India), English does have an increasingly important role on the national cultural stage*: The role and importance of English is evident in the news, entertainment, and other industries* (OC7 India). For one Singaporean respondent, the country is a multiethnic and multilingual nation where a variety of languages are spoken as L1s at home: *we speak a variety of languages as the population comprises of various races and ethnicities* (OC16 Singapore) yet for other Singaporeans English is also an L1 spoken in the home: *I speak English at home and English is my first language* (OC2 Singapore). The same participant perceives English to be widespread: *Singaporeans know english very well, it's spoken everywhere.*

"Personal Impressions of Proficiency"

Participants expressed a mostly high impression of their own English proficiency: *I have a level that is considered high in my country* (OC17 Malawi). A respondent from Pakistan estimated their level at a modest "B2" (OC4 Pakistan), and a similar impression of proficiency was expressed by one respondent from Nigeria: *My level of English could be assessed as 60+%* (OC5 Nigeria). However, none of the 11 respondents studying in the UK claimed to have encountered any specific language problems on their courses at that point and expressed confidence in their proficiency levels: *I would say my level of English was above what I was needed for MA studies* (OC1 Cameroon); *Fluent level* (OC6 Nigeria); *Entirely appropriate* (OC7 India), while others considered whether a typical high school graduate from their country would have a suitable level of English for university study abroad: *he can cope well, without any problem* (OC14 Cameroon). These comments were often linked (a) to their past academic experiences in English, most having been taught in the

medium of English all their lives: *Mode of learning, teaching and speaking is matching those of the UK* (OC11 Malawi) and (b) to the acquisition of previous qualifications in the medium of English: *I already had a post-grad degree in English* (OC3 South Africa).

2. Admission Experiences into UK universities

The second area which was explored using the more open questions was the admission experiences of the respondents, specifically those entering the UK HE system and what role their proficiency and SELTs and CSELTs played. The following themes emerged:

- Affronted
- Reasons for non-exemption from SELTs
- Previous Educational Attainment
- Caveats (to do with language requirements and admissions)
- Quality of the CSELT
- Experiencing a SELT
- Ability to Cope with Academia

"Affronted"

Respondents to the survey generally reacted with objections to their country not being included as a MESC (Majority English-speaking Country), and therefore exempted from doing a SELT: *i feel disowned* (OC5 Nigeria); *Annoyed… It is unfair to insist that students whose entire education has been in English do an English language test* (OC7 India);

Somehow Disrespected (OC10 Cameroon): *bad* (OC8 Hong *Kong); I really feel cheated and disappointed* (OC12 Cameroon)*; Insulted… A SELT is imposed, the prospective student's background and academic credentials justifying his/her English proficiency ignored* (OC13 Cameroon). One respondent was offended that his professional experience had not been taken into consideration: *I felt belittled especially having been an English Language teacher for 10 years at the time* (OC1 Cameroon). A little sarcasm or humour was detected at another student's passing of an IELTS test: *I was then deemed able to survive an unbroken academic study conducted purely in the medium of English* (OC3 South Africa) and one was offended by the content of the SELT they had been asked to take (IELTS); *I felt offended by the test content which I thought did not reflect the demands of academic discourse. The speaking test was a on a subject of gossip* (OC1 Cameroon).

"Reasons for non-exemption from SELTs"

A variety of reasons were given for having to prove English proficiency, one reason was levels of economic development: *Perhaps our being labeled a developing country casts a doubt on the development of our proficiency in English* (OC3 South Africa); *Is it because we are a developing country?* (OC6 Nigeria). Others were more forthright: *I think the UKBA is quite ignorant. The officials there need to broaden their understanding of the world* (OC7 India). Others saw political reasons for the policy (OC8 Hong Kong, OC10 Cameroon, OC13 Cameroon) but did not always elaborate on what they meant by the term "political". Some did do so: *I think this is mainly due to the fact that the English do not have any political or economic interest in that part of Cameroon which they colonised and have happily handed them over to the French* (OC1 Cameroon); *From my experience I think we and everything gravely marginalized by most developed countries* (OC13 Cameroon).

Other respondents, although displaying affronted feelings could see a justification for their compatriots having to take a SELT: *The level of English may not be up to standards. There are serious problems in the standard of English teaching in Malawian schools, such that MSCE holders are not that proficient in English* (OC17 Malawi); *poor ranking on academic index* (OC14 Malawi). The South African respondent saw the validity in the taking of an IAET (IELTS): *While the IELTS tests may be improved, an institution needs to know the level of English proficiency of its potential students. You don't want a situation in which after a month or so, the candidate cannot cope with the medium in which the academic work is imparted, let alone the student being unable to meet the academic tasks given* (OC3 South Africa).

"Previous Educational Attainment"

Many of the respondents already had a high level of education in the medium of English. Sometimes their CSELT or educational attainment in English was valued and recognised in the admissions decisions: *yes it was. I never schooled abroad and did not have to undergo any foreign test to be admitted into University of Leeds. I was happy* (OC5 Nigeria). For others their CSELTs were not recognised or used: *I don't think it was valued at the time of admissions, but once the faculty at the UK university met and interacted with me, they recognised my English language proficiency* (OC7 India); *No. There was no reference to my MSCE qualifications. I was disappointed* (OC17 Malawi).

Within the survey, Likert scales were used to further gauge how respondents felt about having to do a SELT, with the intention of triangulating these with the more open comments

on the same topics. Respondents were asked to indicate to what extent they agreed or disagreed with a statement. These statements were formulated in response to comments on proficiency and CSELTs expressed in the first case study tutorials. Students were asked to what extent they agreed with each statement from "strongly agree" to "strongly disagree". Questions 28, 29 and 32 in the survey dealt with attitudes towards taking SELTs as outer circle students:

28) Students who have had all their education in the language of English should not have to do a SELT

Strongly Agree	12
Agree	3
Disagree	2
Strongly Disagree	0

29) Students from Commonwealth countries should still have to do a SELT because English is not spoken widely enough in those countries

Strongly Agree	2
Agree	4
Disagree	7
Strongly Disagree	4

32) Being taught all their school subjects in the medium of English is better proof of a student's English proficiency than a SELT test score

Strongly Agree	10
Agree	4
Disagree	3
Strongly Disagree	0

The results above show that being educated in the medium of English was a strong factor in students' feelings that they should not have to do a SELT (Q28 & 32). Meanwhile there was some acknowledgement that some individuals in an outer circle country may have to do a SELT due to wider proficiency issues in English (Q29).

"Caveats"

In terms of ascertaining the language proficiency of outer circle students for university admissions, a multitude of caveats and considerations as to the past educational experiences of students were suggested. The high educational attainment in English of the respondents seems to have been a factor in their feelings of being affronted at being asked to take a SELT. Hence, some suggested that if it was the case that someone had been educated in English at degree level in their country of origin then this should allow them to bypass a SELT test: *If already obtained a degree taught in English, then no SELT* (OC6 Nigeria); *Only those that do not have a university qualification should be required to take the tests* (OC17 Malawi).

Others rejected the need to do a SELT at all suggesting that a good score in the CSELT would be sufficient proof for admittance: *Good pass during high school should be a guaratee… SELT test should be abolished in commonwealth countries* (OC14 Malawi); with others adding that a strong performance in a student's subject was indication of sufficient proficiency: *a student with GCE O'Level English and good results at high school or undergraduate level in the pure sciences courses wishing to further studies in related discipline would hardly face considerable difficulties in the UK due to English* (OC15 Cameroon). A compatriot also set conditions for those who should take a SELT: *When overall academic performance is poor with very poor grades in the English language too* (OC10 Cameroon).

Reflecting the concerns expressed above by OC17 (Malawi) about the quality of the CSELTs, others were more specific about grades needed on a CSELT: *I am also aware of the different levels of English proficiency of students from Commonwealth countries. I would be cautious*

and say that if a student from a commonwealth country scored an A grade at the GCE O levels (Cameroon) and has studied in the medium of English they do not need to be subjected to a SELT test (OC1 Cameroon). This was echoed by others: *Depends on their English grade in HKDSE, what their previous degree (s) was/were in and where they were from, and whether they did their secondary education in English* (OC8 – Hong Kong); *Only those students who have not had their education in English should be required to an English test* (OC7 – India).

Others suggested alternative means of ascertaining proficiency in English with the responsibility left to individual institutions: *If student academic back ground is sound in English and Medium of instruction is English in most recent school / college then Interview of student is much better then these tests to access true ability of students* (OC4 Pakistan). Another also left the decisions to the individual university and made comparison with expanding circle students: *when admitted if a student does not measure up to expected standard other measures can apply. The university can work that out. After all, many Chinese students barely speak and write English but they are studying in the UK!!* (OC5 Nigeria).

"Quality of the CSELT"

Some respondents called the quality of a country's CSELT into question as we have seen above in the case of Malawi (OC17). But other respondents valued their education and qualifications in the medium of English as equivalent to the inner circle: *Our teachers go through a UK-like educational system even the Academicians hence the students capable of passing the national exams are of the same footing as UK students of same level… Anglophone students with passes in the High school English language paper, or students with*

B or A grade at the GCE O' Level English are fit for Undergraduate studies (OC10 Cameroon); *Their secondary and high school overall performance can also serve the purpose of a SE*[L]*T since language develops alongside with cognition. Their results from University compulsory courses in English also could serve the purpose of a SE*[L]*T* (OC13 Cameroon)*;A pass in MSCE (English inclusive) is equivalent to international Cambridge certificate exams* (OC14 Malawi). Two of the Likert-scale questions also addressed this issue and indicated that students largely valued their CSELTS with 13 out of 17 respondents disagreeing to question 30 to some extent and 12 out of 17 disagreeing with question 31 to some extent.

Q30) Students from Commonwealth countries should do a SELT because English high/secondary school qualifications from those countries cannot be guaranteed for quality.

Strongly Agree	1
Agree	3
Disagree	7
Strongly Disagree	6

Q31) A SELT test is a better indicator of a student's academic English proficiency than a high/secondary school English qualification from their home country

Strongly Agree	1
Agree	4
Disagree	7
Strongly Disagree	5

"Experiencing a SELT"

For those who took a SELT, there were mixed feelings as to the experience of doing so. A student expressed dissatisfaction with the IELTS scores with time as a factor: *I ran out of time for the writing task, so it was a bit incomplete…* (OC7 India) with a better performance demonstrated in an internal university test: *I scored 100/100 on an English language test*

conducted by the University of Leeds (ibid). Another respondent expressed the view that while his long experience through being educated and qualified in English was not considered, the SELT had a dramatic impact on admissions: *Upon submission of the test scores, my student visa and admission to the university were immediately granted* (OC3 South Africa). However, the same respondent found the test insufficiently specialised: *My gripe about the SELT is that it is generic. It would make sense to set up subject/study-focused tasks which all learners set to follow the study should take* (OC3 South Africa) and another respondent similarly questioned the validity of his IAET in terms of its relevance to academic study: *I did not see any connections between the IELTS test and the language demands of my studies* (OC1 Cameroon).

Two Likert scale questions assessed the value of doing a SELT and revealed ambivalent attitudes towards them. In Q26, the impression given is of students dismissing the SELT as "simply a bureaucratic exercise" (16 out of 17 agreeing), but in Q27 about half the respondents seem to value the SELT as a way to assess their academic English (8 out of 17).

26) Taking a SELT is simply a bureaucratic exercise to get a visa

Strongly Agree	7
Agree	9
Disagree	1
Strongly Disagree	0

27) Taking a SELT is a useful exercise in measuring my ability in academic English

Strongly Agree	0
Agree	8
Disagree	4
Strongly Disagree	5

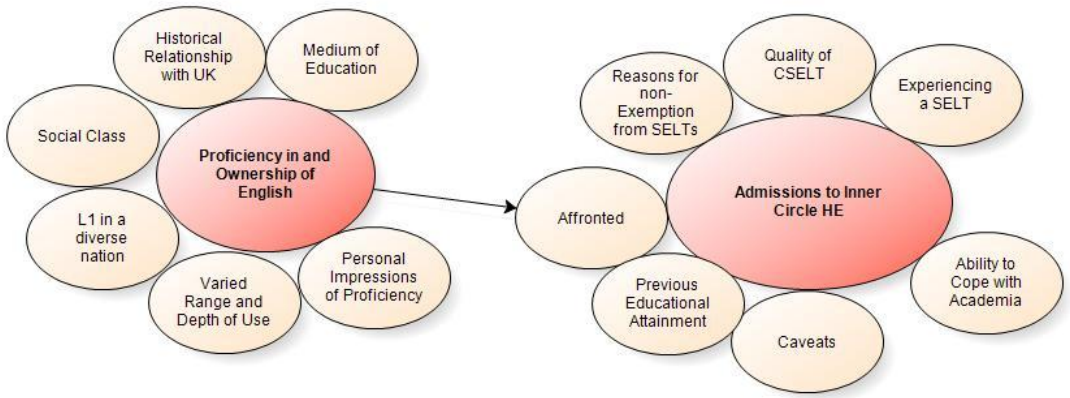
"Ability to Cope with Academia"

Although not investigated in depth, a question relating to academic performance gave a "snap-shot" of views. Very few of the 11 students at university in the UK in the survey indicated they had had any specific problems to do with language at university. There was a comment from a survey participant on her written work: *I remember I received comments on one assignment of some infelicities in my language use* (OC16 Singapore); and from another on encountering varieties of Englishes: *I do face little problem in listening due to difference of English accents in day to day life* (OC4 Pakistan). Another student was quite confident in his academic ability and language: *I completed my MA and PhD and have published in academic journals without any language difficulties* (OC1 Cameroon). Another expressed difficulties relating to the acquisition of a particular skill: *A bit of difficulty because the expectations of writing an essay involves critique which is a bit challenging because you need to read more on different views and then select the view you agree with to further discuss* (OC6 Nigeria). The South African respondent suggested that his compatriots would not have serious issues with academic performance: *The high school exit certificate is comparable to most European states. As a result, most SA students set for academic pursuit can do so without any problem* (OC3 South Africa).

The themes covered above are summarised in the thematic map of the comments below. The two red "bubbles" indicate the two main area of investigation surrounded by the smaller bubbles containing the themes connected to these two areas. The arrow indicates that the themes from the proficiency (and ownership) in English of the participants explain their

"affronted" feelings at not being exempt from SELTs in having to demonstrate their proficiency for admission into UK higher education.

Figure 8: Thematic Map of Outer Circle Survey



4.1.4.2 Interviews with Pearson Country Representatives

RQ3 (PERCEPTIONS OF PROFICIENCY- OWNERSHIP): What are stakeholder perceptions of outer circle proficiency and ownership of English for university entrance?

4.1.4.2.1 Data Collection

Semi-structured interviews were conducted with Pearson country representatives from India, Malaysia and Nigeria. The purpose of the interviews was to collect views on the particular characteristics of outer circle test takers from those countries in terms of:

- Their general performance in the test
- Reasons for doing the PTEA
- Their proficiency in the English language as perceived by the Pearson staff member.
- The standard and reliability of their country-specific English qualifications compared to the PTEA.

The interview schedule can be seen in appendix 16. To obtain participants, I contacted Pearson for the contact details of various Pearson country representatives. Email addresses were provided and I initiated contact and arranged Skype interviews with representatives from the African, South Asian and South East Asian areas of the outer circle - Nigeria, India and Malaysia respectively. Representatives were also sought from Kenya, Ghana and other countries but without success. The interviews were conducted and recorded via Skype (audio and video) for up to an hour in each case. Consent forms and research information sheets were sent out to each participant and returned.

4.1.4.2.2 Analytic Approach

The interviews were recorded (for which consent had been given) and notes taken on the issues discussed. The interviews were then transcribed and coded with the help of Nvivo software to produce a thematic description and "thematic map" for each interviewee. The themes were then compared in the analysis and discussion with the themes from the outer circle student survey and with themes that arose in the case study interviews in the discussion of ownership, linguistic proficiency and academic performance of these outer circle students.

The interviews with three Pearson country sales representatives added another stakeholder view of outer circle student proficiency in English and their test performances. These interviews proved very useful in providing opinions from representatives of the testing organisation under study and the findings are presented below for each representative in turn, firstly with a thematic description of the interview under headings (underlined), followed by a thematic map representing the sub-themes discussed in each interview.

4.1.4.2.3 Findings - Indian Representative Thematic Description

The conversation with the Indian representative centered on the experience of the Indian test taker as regards the PTEA and impressions of their proficiency in English and ownership of the language.

Proficiency and Ownership

From the point of view of the Indian representative, English in India is seen as a second language in which many are competent, but do not regard it as their L1: *most of us prefer to speak, I speak my local language at home so most of us want our kids to learn that too but obviously I have, I have a grip on English.* The impression given was that Indian L1s are much more rooted in Indian life whereas English is used more in schooling: *we still have it in school but daily usage it is not that high when we still have so many languages in India.*

The level of proficiency is then dependent on an individual's schooling experiences - specifically whether English has been used as a medium of instruction. However, there are two main types of qualifications in India, the Indian Certificate of Secondary Education (ICSE), which is more recognised internationally and has more emphasis on English, and the Central Board of Secondary Education (CBSE), which has less emphasis on English. As with other countries there are also a lot of independent international schools which teach in the medium of English and have a relationship with institutions abroad, giving their students an advantage: *cos most of these schools have curriculum, curriculum which are aligned to universities abroad at this point in time, so those schools, maybe those people are most confident*. It was stated by the representative that a lot of the students applying to go abroad had already been through an English-medium high school system and, in addition, a tertiary

level English medium education in India. For example, because a lot of the PTEA test takers in India are pursuing post-graduate study abroad, most of them study a first degree in India in English: *obviously these people have more chances of going abroad so they would ideally want to plan to go abroad they would at least study in an English-medium college or university*. In this way, acquisition of academic English may be based on aspiration and migration: *it's spreading out, it's spreading out now it's becoming a dream for everybody to go abroad and settle down:* According to the representative, Indian test takers have acquired a certain proficiency and ownership of English from their education and desire to migrate rather than their actual social and family life. In this way, the representative sees Indian students as confident in reading and writing as opposed to having to interact in speaking and listening: *they're much confident because the, it's only them involved so I guess it comes, you understand, the writing doesn't involve – you don't have to converse lessons so I think it's much easier for them.*

Two related themes were the status of the Indian CSELTs and admission to academia. There was a past relationship and "flexibility" in language requirements between British and Indian institutions with an associated understanding that Indian CSELT qualifications could be trusted and guarantee a level of English, but the representative mentioned that this had recently showed signs of changing: *UK was the one which is very, I mean, they were willing to accept the ICSE ones, the domestic included and the CBSE ones but other countries don't at this point in time and the UK is also getting strict so…*

As expressed by this one representative, the impression he gave of proficiency and ownership of English in India is that the language is seen as a practical second language with most Indians still using their own non-English L1s for domestic life. Academically, the average

Indian student is seen as competent in English from their previous education, particularly in reading and writing. However the representative gave the view that Indian CSELT is becoming less recognised in the context of proof of English for admission to international universities.

Experiencing the PTEA

The representative had not detected any feelings among Indian test takers of objecting to, or being affronted at having to do an IAET. Rather, the test is seen as a purely practical obstacle in their journey abroad: *everyone looks as a secure test as an option to jump I mean maybe a hurdle… I mean, they just have to do it for their visa*. In terms of proficiency as evidenced by test performance, the representative did not rate the average Indian as a notably high scorer: *I mean between B1 and B2 I guess I mean sometimes between there I don't think it's even [inaudible], so that should be the average err…capability of an Indian student.* Preparation for the test is short: *we have a lot of last-minute takers… I've not seen any training go beyond ten days that's how it works yeah.* Test preparation is largely about the key task of familiarising oneself with the format and content of the PTEA rather than improving proficiency as the consequences of not familiarising oneself with the format were expressed by the representative: *but if you don't know the format and go for the examination preparing for competition tests I'm sure it's going to be difficult*.

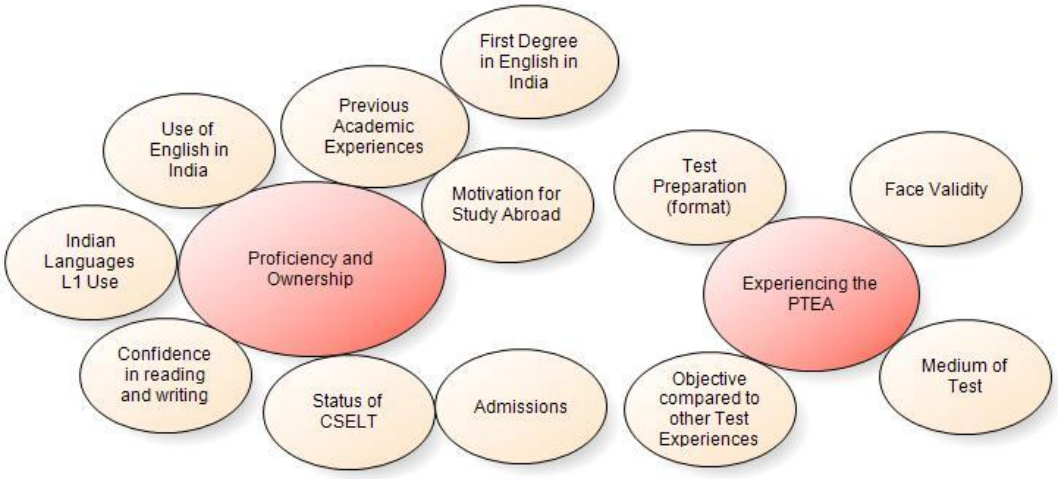
The test has an appeal in terms of its "face validity" from being computer based as the representative said that the average Indian test taker is worried about being judged in the speaking part of the test: *the biggest challenge with the Indian student today is err…err, err, the fear of facing an examiner – subjectivity – I think we're being very very objective in our*

examinations I think that's been a win for us. This is something which, he says, has affected the test performance of Indians on competitor exams: *people taking IELTS two, four, three, three-four times, the fear of the examiner brings the score down*. The PTEA is seen as valid given its purpose to measure whether a student will be able to cope at university: *I think it's accent-neutral so [inaudible] it's only the academic part of it we're looking at - whether you are able to understand a lecture, whether you are able to converse…*. However, the test is still seen to favour an urban class of test takers with familiarity with and access to the medium of the test in a technological sense: *there are cities, obviously the metros, people, it's not challenging obviously… everybody's on Facebook and everything but if you go to…maybe a remote place yes there is a challenge at this point in time they are not too happy going on a computer*.

The impression given by the representative regarding experiencing the PTEA is that most Indian test-takers engage in minimal preparation for the test and score moderately well on it. The test is seen as more objective because it is administered via computer rather than involving a more subjective human element (from the point of view of the Indian test taker). According to the representative, many Indians have had their education in the medium of English to undergraduate level but do not object to doing an IAET. Their motivation for study abroad is very high, linked as it is with a desire to migrate. Their confidence in reading and written English is high and their preparation for the test is undertaken in a purely practical sense to familiarise themselves with the format of the test as a barrier to be overcome in pursuit of the goal of going overseas to study. Although some Indians may not like the computer medium of the test, the PTEA is seen to be more objective (and have face validity) in comparison to experiences of other tests (such as IELTs) because the factor of a human

examiner is removed. The thematic map below summarises these descriptions and themes above.

Figure 9: Thematic Map of Interview with Indian Representative



4.1.4.2.4 Findings - Malaysian Representative Thematic Description

The themes emerging from the interview with the Malaysian representative centered around the areas of the proficiency and ownership of English, experiencing the PTEA and admissions issues.

Proficiency in, and Ownership of English

The wider theme of proficiency and ownership comprises the sub-themes such as "English in Malaysia" and "English Proficiency". As with India, the status of English in Malaysia is complex because of the history of English in Malaysia and its use among different ethnic and social groups in the population (Hashim, 2014). The national language for communication is Malay but much of the population is bilingual or trilingual in English, Malay and other languages (typically Tamil and Chinese languages). According to the representative, most people have a "good standard" of English. In common with many outer circle

Commonwealth countries, the education system includes school qualifications equivalent to the former UK 'O' Level inherited from colonial days such as the Sijil Pelajaran Malaysia (SPM) a Malaysian certificate of education: *Malaysian education system is largely based from a UK education system…right so SPM is 'O' levels.* Regarding the official medium of education, the representative mentioned that the majority of Malaysian schools now teach through the medium of Malay since the language policy shift from English to Malay in the 1970s which has had an effect on the general language proficiency in the country: *Yeah in the late seventies we switched from English to Malay, I think there was a bit of a national agenda to promote more of Malaysian, err Malay language so English, English got little bit diluted then but still reasonably well-spoken.* The country also has Chinese and English medium schools. The latter tend to be private, international schools which arose from the switch to Malay and favour a certain social class: *international schools has come up to cater for even for middle-class*, and: *a lot of Malaysian students joined private colleges, in Malaysia it's all taught in English*.

In contrast to the Indian Pearson representative, the Malaysian representative mentioned some test takers' feelings of being affronted at being asked to take a test such as the PTEA: *when the students feels they have to sit for, to prove their English they get a little bit frustrated, yeah right 'why why do I have to sit for it, my English is good'*. However, most accept the need to do it to obtain a visa as part of the bureaucracy of studying overseas. In terms of a level of "English Proficiency" the representative stated that the PTEA scores would be higher than the world average though he added: *with Malaysians I think they're not at Singapore level* and in terms of mother tongue speakers he guessed that only around 10% of the population would regard English as an L1 though he said that it is increasing,

particularly in urban middle-class areas (this compares to 30.4% of the test-taking sample from Malaysia stating English as an L1 in the 2014 PTEA data above).

Admissions Issues

This proficiency in English then links to a complex admissions situation where for many years UK universities have accepted Malaysian students based on the strength of their CSELTs of which there are two, the SPM (generally at GCSE level) and a higher paper called the "1119" which lies somewhere between GCSE and A-level. In the past, most UK universities have been happy to accept Malaysians based on these qualifications and an academic "track-record" of students from a particular college. Before students from the P.R. China arrived in the UK in large numbers to pursue higher education, Malaysians comprised the largest cohort of international students at UK universities, reflecting the long relations between UK institutions and Malaysian colleges: *before 1997 Malaysian students used to be the number one, number one international students in the UK*. This situation included a level of "comfort" and "trust" characterising the long relationship between UK universities and Malaysian colleges: *a lot of UK universities are quite familiar with Malaysian education … there is a lot of comfort level between a lot of UK universities and Malaysian err institutions.* This has led to a flexibility in admissions requirements: *I think quite substantial number might not need to show their English language, language proficiency, so that means they don't need to sit for PTE or…or for IELTS or TOEFL*.

This flexibility has, however, begun to change and is manifested in an increasing strictness in admissions where some universities no longer accept the CSELTs and instead are now insisting on SELTs or other IAETs as proof of English proficiency. However there is still a

commercial advantage to being more flexible with admissions policies and many UK universities, especially the former polytechnics, are remaining flexible with Malaysian students by not insisting on a SELT and waiving English requirements: *To put it into advantage for the students to choose that right so they'll tend to waive the English requirement….* This flexibility is based on their past relationship with Malaysian colleges with whom they often have collaborative "top-up" degrees and study programmes already in place: *they they receive a large number of Malaysian students who will go, only do nine months, basically they'll do just one final year, (*Interviewer*: "okay") right so which we call err… 'top-up degree', err sometimes it's called twinning degree right okay…mostly it's with collaboration with the local college in Malaysia*.

However, being flexible on admissions and language requirements as a way of competing with other universities is seen as a risky exercise: *well another thirty to forty percent of universities have a balance of both, they like to be competitive but at the same time they do not want to get into trouble.* This is in reference to balancing being flexible with English language requirements and also considering UKVI policy on SELTs and Tier 4 so as to avoid getting into trouble with the UKVI (according to the representative).

The PTEA itself is seen as a good measure of English and a comprehensive test of skills: *it's comprehensively evaluates your level of English* and together with other SELTS is preferable to a CSELT in order to make an admissions decision: *it helps for the universities to make a proper decision.* In his view this is because an IAET is seen as a reliable measure of English proficiency: *at least there's a barometer to check you, you know.* In contrast, the Malaysian CSELT is not seen as being as comprehensive a test because of the absence of speaking and listening components: *it's no way it evaluates the four communication skills because there's*

no, there's no oral I mean there's no speaking part, there's no listening part, it's just a written test. The representative was worried about the ethics of accepting students with too flexible an admissions policy due to the financial importance of international students: *there's always err not so comfortable feeling with some of what the UK universities does.*
Following on from admissions and into academia, the Malaysian representative claimed that Malaysians tended to do well academically: *generally, Malaysian students have a good track-record academically*. The representative emphasised his preference for Malaysians to sit a SELT (not just PTEA) because then the student and university could be sure that the student would have the requisite language skills in order to cope at university. Otherwise he foresaw a situation whereby: *you have to go to UK and have to survive in an English-speaking environment and then do an academic qualification, they might not ended up doing well.*

Experience of the PTEA

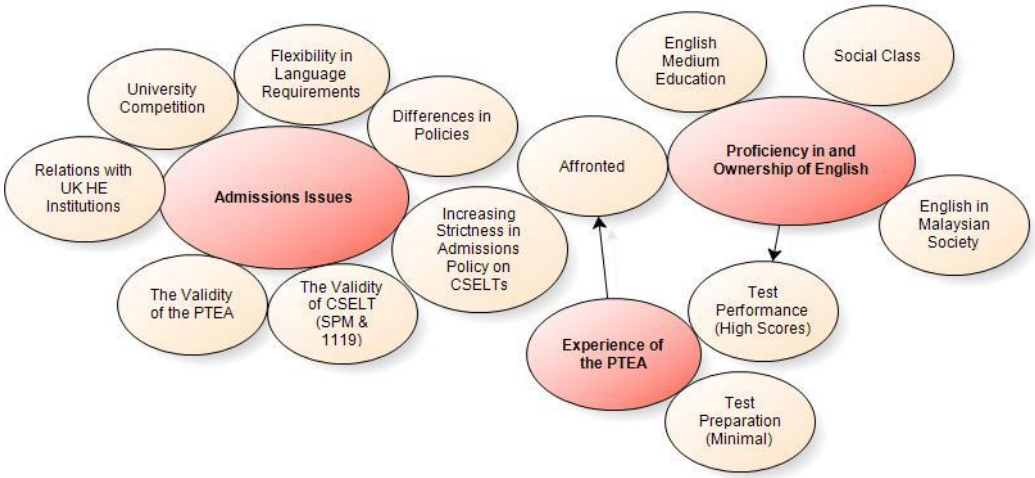
As with the Indian representative, it was stressed that a good test performance should be based on knowledge of the format of the test. The representative mentioned that provision for this is improving in Malaysia in that Pearson are providing training courses for familiarisation with the PTEA. However, the representative implied that test preparation is still not ideal with often minimal consideration given by students to preparing for the test: *generally still a, students who sitting it, who have to sit for it are getting, are go and do it and hope and… that sort of thing yeah.* In terms of test performance it was felt that Malaysians would generally score higher than the global average on the PTEA: *I think it's about seventy per cent of the students are scoring more than 50* although he felt that some would struggle in any SELT if they had not had English as a medium of instruction, as is the case in the Malay-

medium schools. Thus, the representative felt that most Malaysians would regard the PTEA as difficult: *the general perception is erm…not many students will say it's an easy test*.

From what the representative said, it can be concluded that the Malaysians' sense of ownership of and proficiency in English depends on their cultural background and class background, and in common with what the Indian representative said, dependent to a large extent on whether they have been sent to an English-medium school. In general, English proficiency has been declining since the switch to Malay-only instruction in state schools. However, as with India, a large middle class has the means to by-pass this policy and educate their children entirely in English. Due to the long historical relationship between the UK and Malaysia in the realm of education, many UK HE institutions waive the language requirements of students depending on the specific arrangements between individual British universities and Malaysian colleges. However, as with Indian CSELTs, this is changing in line with admissions and language policy changes in the UK, although Malaysian students are still regarded as high achievers academically.

The thematic map below illustrates the opinions of the representative, that the English proficiency of the Malaysian student is dependent on their backgrounds in learning and using English in their school systems. It means that their CSELTs are often still accepted in place of SELTs especially for those universities with whom the Malaysian colleges have a close and long relationship. This relationship also explains the flexibility in the admissions system regarding proof of English proficiency. Their generally high proficiency is reflected in higher than average PTEA scores (as claimed by the representative and indicated in the 2014 PTEA data as above).

Figure 10: Thematic Map of Interview with Malaysian Representative



4.1.4.2.5 Findings - Nigerian Representative Thematic Description

The interview with the Nigerian representative produced themes around the areas of proficiency in and ownership of English, experiencing the PTEA and its Validity, and Proficiency for Academia.

<u>Proficiency in, and Ownership of English</u>

The theme of the proficiency and ownership of English is a lot more apparent in the Pearson Nigerian representative interview compared with the Malaysian and Indian interviews. A number of comments were made about the nature of "English in Nigeria", which suggests that English is deeply embedded in Nigerian society, both in the state and in the homes of Nigerians to a larger degree than in India or Malaysia. Nigeria is a highly multi-ethnic and multi-lingual nation which means that English is used widely as a means of communication among Nigerians in towns and cities where people will not readily be able to communicate without English as a Lingua Franca or Second Language: *in the streets of Lagos or Abuja you aren't going to speak your dialect … your mother-tongue… so the next person, and the next person, isn't from your village … so, so it's also English that they use….* Furthermore, for

many younger Nigerians, the representative mentioned that English is increasingly an L1 in the home: *most young people erm…may not have their Nigerian languages as L1 cos they start from day one, their parents even speak English in their homes*. This suggests that for many Nigerians, English is not only a second language but a *de facto* first language which matches comments made in the outer circle survey. Furthermore, part of the nature of English in Nigeria lies in the education system reflected in "Previous Academic Experience" that suggests ownership of English through Nigerians having been taught in the medium of English. Indeed, English medium education is the norm: *it is a language of education.* This institutional use of English starts from nursery school onwards – again - because of the multi-lingual nature of Nigerian towns and cities: *as low as the, the pre-primary school, pre-primary, English is used because err…people from different tribes, and linguistic backgrounds, are brought together in, in Nigeria, in any city so it will be difficult for you to use the L1, the mother tongue.*

Looking at "English Proficiency" in relationship to ownership of the language, the representative suggested that English in Nigeria is part of the country's linguistic identity: *it's not a foreign language for Nigerians.* That's not to say that all Nigerians are equally proficient in English by any means but there is a suggestion that the proficiency of school-leavers - the relevant group for admission to overseas universities - is already at a high level and that this is linked to their PTEA performances and experiences in the Nigerian school system: *people who just come out of, or from school, take the test do fantastically well, and better, their scores are higher and better …it's a reinforcement of what they've already doing in school.* However, the representative suggests that even though many Nigerians possess a high level of English proficiency, there is still a need to demonstrate it*: you might even be better than even the first language speaker but you need to show it.* This relates to his later

comments (see below) about the validity of the CSELT (good in testing reading and writing skills but deficient in terms of testing speaking and listening skills). In terms of proficiency in separate skills, although acknowledging that spoken English in Nigeria would be quite distinctive and influenced by the various other local L1s, performance in the other skills is not regarded as a problem by the representative: *if it is writing, no problem, if it is reading no problem.* Regarding the specific area of spoken language proficiency where there may be communicative problems, the Nigerian representative suggests that accent could be an issue in terms of being understood internationally: *because it's a second language the whole idea is to teach them to the level of international intelligibility.*

The sub-theme of "affronted" also contributes towards the theme of ownership specifically regarding the reactions of specific stakeholders – teachers and parents – to the students having to take an exam such as the PTEA in order to demonstrate their English proficiency and secure a visa for the UK. The students themselves do not raise objections to being asked to take the test but the representative mentions that parents and teachers have occasionally done so: *I haven't seen anybody saying 'why should we take the test' … but I have seen a few adults, erm…especially those who are teachers of English, … saying 'why, why do our students, err…do this test?* This phenomenon was mentioned a few times in the data referring to the representative's experiences at conferences in Nigeria. The representative himself wondered aloud whether, if he was in a higher official position, he too would question the obligation for students to take tests such as PTEA: *if I were in government in Nigeria, I could make a case and say 'no, I don't want students to take this test' because we speak English as a Second Language.* This aspect was not mentioned by the Malaysian or Indian representatives.

Experiencing the PTEA and its validity

Crossing over from proficiency and ownership to the area of experiencing the PTEA, contributing to the sub-themes of "Test Preparation" and "Test Performance", the representative regards the English proficiency of the students as of an ESL nature and therefore the type of preparation they undertake for the PTEA as not the same as in a country where English is less commonly spoken: *the preparation in Nigeria is not like doing a course in India, no, we don't do language courses here… you don't need to start teaching them English…* (However, the Indian Pearson representative made the same statement in that Indian test takers only prepare for the format of the exam). Similarly, what the Pearson Nigerian representative describes regarding the test preparation courses suggests that the main concern is to familiarise the students with the format of the test: *what we do here is to just, give them preparation… [inaudible] you know, tutorials on a particular test, this particular test … this is the pattern of the test, this is what to expect, this is, for listening … this is what you do…give them tips about this test … answers, the questions and the tips for the answers, and things like that … that's what we do in Nigeria.* The obligation of having to do the test does not engender an increase in language study but rather a need to know the format of the test and to familiarise oneself with the medium of the test.

In countries such as Nigeria where access to computers is not universal, the medium of the test may also be an issue in the experience of the PTEA: *because it's a computer-based test so most people who don't do well, isn't because they don't have the skills, for English, well I think that most of them are coming from the hinterland where they haven't been exposed to the use of computers, the majority part of Nigeria isn't still exposed to computers*. As mentioned above, in terms of test performance, a strong performance in the PTEA is very

much seen as the experience of recent school-leavers and as with the Malaysian representative, a supposition is made that Nigerians score higher on the PTEA than the world average: *Nigerian average score is higher than the world average essentially yeah*.

In terms of validity, in common with the Malaysian representative, the Nigerian representative regards the PTEA as more comprehensive than the CSELT, the national Nigerian English exam known as the WAEC (West African Examination Council). He says of the PTEA: *it's balanced, it's for the, it tests everything actually.* This is in contrast to the WAEC which in common with comments from the Malaysian representative on their CSELT, is seen as neglecting the oral/aural components: *…we did not test listening and the, the WAEC, does not also test speaking*. The PTEA is seen as a way of proving your proficiency even if you already possess proficiency from your schooling and WAEC.

However it is significant that the representative himself calls into question the validity of IAETs like PTEA for Anglophone West Africans and suggests that tests may be more valid for Francophone West Africans: *the danger what I, what I am thinking is that, erm…since, because it is the same test that the person who speaks English as a foreign language, takes, so it's more a model for a French-speaking person – a French-speaking West African, much more than a Ghanaian, a Nigerian, a Liberian, a Sierra Leonean*. Despite this, as mentioned above, the representative did not detect any feelings among students of being affronted at being asked to take the test. However, the impression given from the interview is that although the PTEA is regarded as more comprehensive than the local English qualification, more questions are raised as to its validity and suitability for Nigerian test takers who tend to have English as a second or *de facto* first language.

Admissions Issues

An important theme explored in this interview was the link between the validity of the PTEA and the admissions experiences of students to universities in the UK. Firstly, there were conflicting perspectives coming from the Pearson representative both defending and questioning the use of the CSELT and the PTEA for admissions respectively. On the one hand, the CSELT was seen as strongly assessing the key academic skills of reading and writing: *two major language skills are really, really broadly tested and you cannot doubt, the scores of people in those two, two major skills - the skill of reading and the skill of writing*. The authority behind the test, the West African Educational Council, values reading and writing skills as key academic skills (see below). Further validating the CSELT is the fact that the qualification is accepted by most UK universities as sufficient proof of English proficiency: *very many UK universities accept the, West African Examination Council's certificate – school certificate as enough qualification for the students to come over.*

However, he went on to state of the CSELT that*: I think it's testing fifty per cent of the academic English* because, as mentioned previously above, speaking and listening skills are not adequately tested and the qualification in this regard is "*defective"*. Crucially he said that this may mean that: *you do not know whether the student can speak, intelligibly* and hence the need for Nigerian students to take an IAET such as the PTEA. However, these comments are still balanced by his previous critique of the PTEA (see above) and his view that the PTEA may only be regarded as a "reinforcement" of students' *existing* English proficiency for university study.

Proficiency for Academia

The last theme emerging from the interview was the one of actual linguistic performance in academia - references made to linguistic skills and academic skills. The authorities behind the CSELT itself see reading and writing as fundamental skills for academia: *what they believe strongly that written English gives what level you are and then reading, because that's what you are going to be facing, in the academic environment.* This is already reflected in the CSELT which focuses on assessing these skills and assumes that at university, students*: are going to be reading and that means you are going to be translating, you are going to be interpreting, you are going to be extrapolating.* In terms of rigorously testing the key academic skills of reading and writing, the Nigerian qualification appears valid according to the representative: *huge writing exercises, huge reading exercises, comprehension exercises, summary [inaudible] exercises, you know, creative writing, and things like that, grammar, then smaller ones, enabling skills like grammar, vocabulary, development and things like that, those are hugely, hugely tested.* This suggests that students who pass their CSELT at a good level will theoretically cope with the linguistic rigours of academic reading and writing at university.

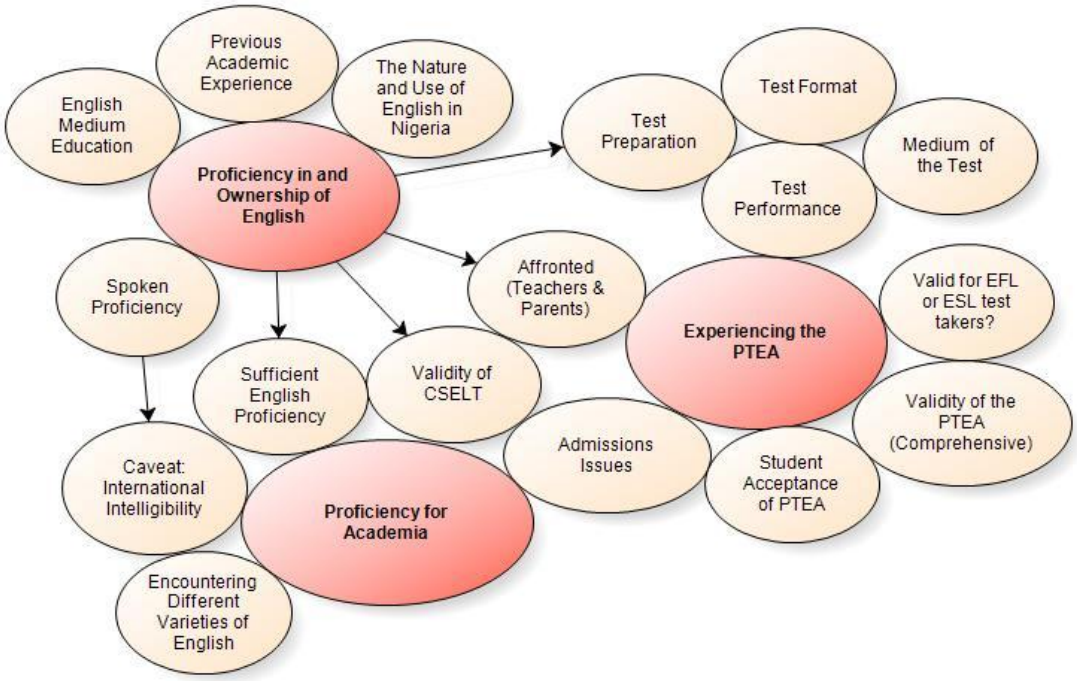
Speaking and listening skills are less valued by the authorities behind the CSELT because of English dialectal differences across borders in West Africa where the test is administered: *Their idea is that after all, that would be, there are dialectical differences erm, even within the English-speaking people* but the representative suggests that there is a need to test whether the students can communicate and are "internationally intelligible". This is related to the sub-theme of "Encountering Different Varieties" of English at university, and this aspect is anticipated as possibly affecting academic performance: *different people who might be*

lecturers, have different accents (Inteviewer: "yes, interesting") so sometimes it's not easy to, to, to decipher the accent of say an Ind…like it's difficult for a Nigerian to actually easily grasp with err…an Indian accent. Lack of assessment of these two oral and aural communicative skills is conceded as a problem for the CSELT in Nigeria. The representative suggests that language at university will generally not be a problem for Nigerian students but the idea that the West African variety may pose some challenges in communication is apparent and so he sees the preparation of Nigerian students for study abroad as needing to ensure they are intelligible to other English speakers. However, this aspect is not considered to be a major reason why Nigerian or outer circle students would be significantly disadvantaged in academia. Indeed, the representative later states that this problem may be less significant than previously thought: *usually everybody carries some element of his L1 – by L1 I mean the first language …it's comprehensible but erm you will still see the accent you still hear the accent.*

Overall, the evidence from the Pearson Nigeria interview is that Nigerian test takers are linguistically prepared for university through their CSELT and home education system with some caveats about adjusting to being internationally intelligible and encountering other varieties of Englishes. In urban, middle-class Nigerian society, English is an important part of the lives of Nigerians, and is often used as a *de facto* L1 despite the presence of another L1. Therefore, Nigerian students are theoretically linguistically prepared for communication at university. In terms of test validity, the PTEA is regarded as a more comprehensive measure of language proficiency than their CSELT – the WAEC, particularly with regards the spoken and listening components of the test but it is unclear as to what extent the PTEA assesses reading or writing skills any better than the WAEC. According to the representative, there are indications of some feelings of being affronted from some stakeholders as to why Nigerians

have to take the PTEA test but from students themselves there is more of an acceptance of the requirement. It is claimed that Nigerian test-takers score above PTEA world averages and prepare for the test by largely familiarising themselves with the format. The medium of the test may pose some difficulties for citizens who are less familiar with computers particularly in rural contexts. The thematic map below illustrates this summary of the interview with the Nigerian representative. The arrows in the map indicate that the ownership of English of Nigerian test takers comes across very strongly in the interview and impacts on the Nigerian experience of the PTEA and how Nigerians demonstrate their proficiency for admission to university. Some stakeholders express some objections to doing a SELT, represented by the theme "affronted". The CSELT is an expression of English language proficiency and ownership of English from the domestic education system and demonstrates sufficient English proficiency for university.

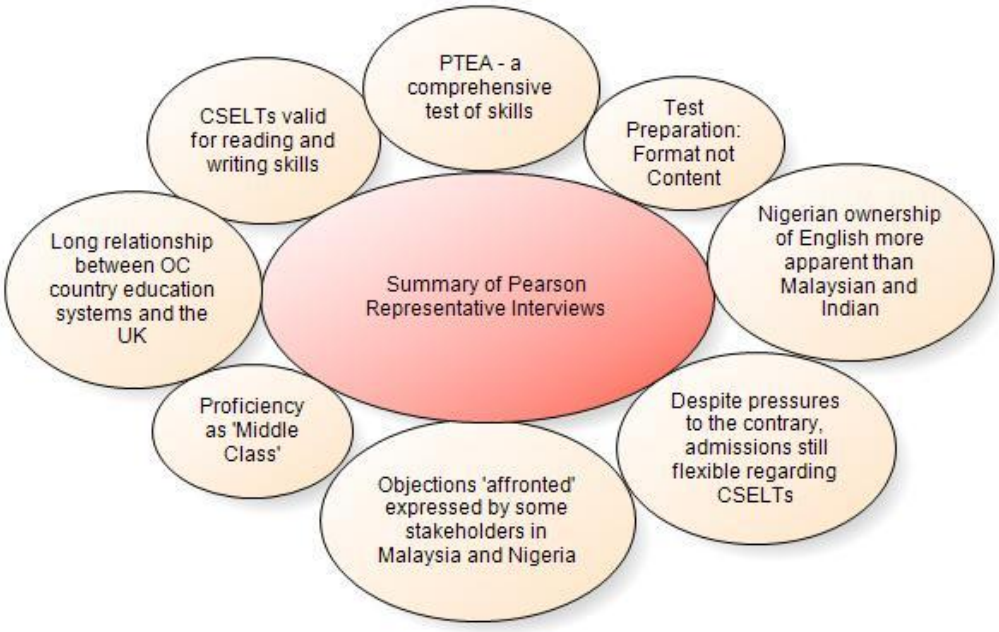
Figure 11: Thematic Map of Interview with Nigerian Representative



4.1.4.2.6 Findings – Summary of Pearson Interviews

The three interviews together illustrate the complex nature of the range and depth of English usage depending on the particular outer circle society. However, the three interviews have themes in common which are represented in the thematic map below and shall be explored more in the discussion section of this thesis.

Figure 12: Summary of all Pearson Representative Interviews



4.2 Strand 2: Predictive Validity and Academic Performance – The Case Study

Strand 2 Case Study Participants

- 47 individuals answered the recruitment survey (of whom 46 provided useable data) and their data was used to trial the self-rating tool, and to gauge opinions and experiences of the PTEA as well as to recruit individuals for the case study.

- Four individuals (3 Nigerians and 1 Ghanaian) agreed to a follow-up tutorial after being contacted through the response to the recruitment survey. These four individuals became the four "cases" in the case study for the study on the predictive validity of the PTEA and academic performance.

The research questions addressed by the case study are 4 and 5 below:

RQ4 (PREDICTIVE VALIDITY): How do the PTEA score profiles of individual students compare to their actual language behaviour and academic performance in their first year?

RQ5 (VARIABLES): What are the multiple variables that affect the academic performance of outer circle test takers on their chosen courses?

4.2.1 The Case Study Recruitment Survey

A recruitment survey was first carried out mainly to find individuals who were studying at university in the UK to take part in the case study.

4.2.1.1 Data Collection

The first stage of the data collection used an online survey (appendix 7) distributed via an email sent through the Pearson organisation to previous test takers of the PTEA who had given their permission to be contacted for research purposes. The purpose of the survey was to capture a variety of nationalities from the outer and expanding circles in order to:

- Recruit participants for the case study
- Trial the self-rating tool of proficiency to be used in the case study
- Gather some initial opinions on test performance and perspectives on proficiency
- Identify any initial emerging themes from the comments

The survey consisted of closed questions to collect biographical data, a tool for self-rating proficiency, followed by open-ended questions and a Likert scale asking for opinions on participants' test scores and their test taking experience. The survey was put online because it could be distributed easily via email to a variety of potential respondents in institutions in the UK using email lists such as that of the British Association for Applied Linguistics (BAAL), and email lists from Pearson. The survey was accessed via a hyperlink in an email and could be completed in 20 minutes and the results were easily collated from the university website which hosted the survey (as outlined in the methodology chapter above).

The email was sent out in June 2013 and again in October 2013. The June survey was intended for students who had started courses in February 2013 (for example those on diploma courses leading to a Masters or students on pre-sessional courses at the universities) and the second wave of questionnaires was sent out in October to survey students who had started their courses in September/October 2013. Students self-reported their PTEA scores

including both the communicative and enabling skills scores. If the students agreed to be part of the case study their self-reported scores would be verified by asking for their official transcript of results in the first tutorials.

In effect, the first mailing list in June 2013 acted as the pilot survey. Based on the piloting results, I amended only two questions in terms of clarity of wording in the questions. Students reported their PTEA score profile and had an opportunity to make initial comments about their test performance and proficiency. These answers were followed up on subsequently in the first tutorials with the participants who agreed to further participation in the case study.

4.2.1.2 Analytic Approach

Quantitative data in the form of PTEA scores were collected but because they were self-reported and constituted a very small sample in comparison to the 2014 PTEA data set, they were not analysed in terms of mean scores or to distinguish outer circle performance on the test from expanding circle performance. Consequently, they are therefore not reported here. Self-reported PTEA scores were only used to trial the self-rating tool for proficiency. The self-rating choices of the entire sample were tabulated to see to what extent the self-rating matched the self-reported PTEA scores given by the participants using the CEFR as point of comparion. The qualitative comments were collected, coded and analysed for initial themes that emerged on test performance and perceptions of proficiency and informed the content of the initial tutorials in strand 2, for example regarding the cases' experiences of taking the PTEA which were probed in the initial tutorials.

4.2.1.3 Findings - The Recruitment Sample Descriptive Statistics

There were 47 respondents to the call for participants, which was disappointing but not uncommon in research of this type (Kerstjens & Nery, 2000; Hill et al, 1999). This was despite best efforts to disseminate the surveys through email, social networks and social media and A4-posters on university notice boards. Two of the respondents provided incomplete data for test scores so only 45 participants' data was used. The summary of the respondents is below:

Demographic Data

- The useable data in the recruitment survey was for 45 individuals
- The age range was 30 years from 17-46, mean age 29.13, SD 7.585 (n=45)
- 33.33 % respondents were female and 66.66% were male.

Nationality

- 2 test takers were from 2 inner circle countries (Australia and Canada)
- 31 test takers were from 7 outer circle countries (Nigeria, India, Pakistan, Kenya, Cameroon, Ghana and South Africa)
- 12 test takers were from 11 expanding circle countries (Turkey, Brazil, Colombia, Denmark, Iraq, Israel, Japan, Kuwait, Libya, Romania and South Sudan).

Table 59: Recruitment Survey Participants (n=45) by World Region

Region	Africa	Europe	The Americas	The Middle East	South Asia	East & Southeast Asia	Oceania
No. of Participants	19	2	3	5	14	1	1

In terms of regions dominating the sample, there was a preponderance of African and South Asian test takers (table 59) and within that there were 12 participants from Nigeria and 11

from India. This matches the pattern in the large PTEA data set in the first strand of the research where India and Nigeria are represented by at least 500 test takers each. A notable absence in the recruitment survey sample is East & Southeast Asia where there were no respondents to the case study recruitment survey from Malaysia or China, despite large samples from both countries (n=500) being recorded in the PTEA test data for 2014.

In terms of the language background of the participants, a whole range of first languages was in evidence, mainly a variety of African and South Asian languages. It was a multilingual sample as every respondent apart from one, (Australian), indicated knowledge of another language. Only three individuals identified English as their first language (2 Nigerian, 1 Australian).

4.2.1.4 Findings – The Recruitment Sample Self-rating of Proficiency

All students were asked to rate their own language proficiency using the descriptors provided. This would test the tool used to provide different perspectives on proficiency for the case study. The results are displayed in the summary tables below, and the five tables in appendix 17. In tables 60 and 61 there is a summary of the information from the more detailed tables in the appendices. These tables illustrate the "match" between a test taker's PTEA scores and their self-rating of proficiency with the CEFR as a comparison tool.

Table 60: Proportion of Students Over and Under Rating their PTEA/CEFR Levels

	Overall Scores+	Listening	Reading	Speaking	Writing
Participants overestimating	27/46 58.69%	21/45 46.66%	28/45 62.22%	11/45 24.44%	17/45 37.77%
Participants matching levels	16/46 34.78%	18/45 40%	11/45 24.44%	23/45 51.11%	14/45 31.11%
Participants underestimating	3/46 6.52%	6/45 13.33%	6/45 13.33%	11/45 24.44%	14/45 31.11%

+ A Russian participant gave his overall scores but no communicative scores were given, hence n=46 for overall

It can be seen that most participants appeared to over-rate themselves for their overall and reading scores in comparison to their self-reported PTEA scores (58.56% and 62.22% respectively). This indicates at the very least a disagreement or mismatch between the two measures of proficiency. The skill with the largest proportion of the sample matching their PTEA and self-rating was in speaking with around half the sample (51.11%). A third of the sample (31.11%) underestimated their level in writing but 37.7% of the sample overestimated their written proficiency using the self-rating tool. There are high percentages of the sample overestimating their proficiency levels in comparison to PTEA test scores, with the exception of the productive skills of speaking (24.44%) and writing (37.77%).

Table 61 Self-Rated CEFR Levels vs. Actual PTEA CEFR Level

	Number of students rating themselves at this level in each skill vs. Numbers who actually Scored at that level (in parentheses)				
Level (CEFR)	Overall (n=46)	Listening (n=45)	Reading (n=45)	Speaking (n=45)	Writing (n=45)
C1	26 (10)	24 (12)	15 (6)	15 (14)	10 (10)
B2	15 (15)	12 (15)	7 (15)	13 (14)	16 (17)
B1+	3 (10)	4 (6)	19 (5)	8 (5)	9 (5)
B1	1 (7)	3 (9)	2 (5)	6 (8)	5 (8)
A2	1 (4)	2 (2)	2 (10)	3 (2)	5 (2)
A1	0 (0)	0 (1)	0 (4)	0 (2)	0 (3)

Table 61 summarises the more detailed tables in the appendices (appendix 17). It shows the number of students who self-rated themselves at each CEFR level overall and in each communicative skill. The numbers in parentheses (n) indicate the actual number of test takers who scored at that level (self-reported). It can be seen, for example, that only 5 students self-rated their overall proficiency at below the B2 university threshold level, whereas 21 test takers actually scored below B2 level overall. Elsewhere, however, in speaking and writing, there seemed to be a match between the numbers at B2 and C1 levels.

The data from tables 60 and 61 and in appendix 17 is useful in flagging up a possible phenomenon with a self-rating tool of this kind – namely that individuals have a tendency to over-rate their own skill levels (Dunning & Kruger, 2006).

4.2.1.5 Findings – The Recruitment Sample Thematic Description

The recruitment survey invited comments on three areas:

1. Preparation for the PTEA
2. Experiencing the PTEA
3. Use of the PTEA in admissions

Comments are reproduced below exactly as they were typed into the survey by the participants themselves, therefore the transcripts have spelling errors in them.

<u>1. Preparation for the PTEA</u>

Four themes came out of the broader preparation theme: a lack of preparation; preparing for the format of the test; preparing in terms of English language and preparing for a specific discipline or subject area. These are illustrated one by one below:

Lack of preparation

Respondents expressed a lack of preparation for the PTEA, perhaps due to over-confidence: *For the test itself I did not prepare at all. On the day before I had a little practice on the*

material I had from them, by reading and answering some of the questions. I thought it would be much easier and I was very confident I would do well (Student #39). Others gave the impression that they had just turned up on the day: *All i just did was wait for the day, walk into the testing center then i did the best i can* (Student #42).

Preparation for the PTEA format

One step up from no preparation at all was a recognition of the need to be familiar with at least the format and techniques needed in the test: *Preparing to attend exam preparation classes to enrich my skills on how best to approach and pass exam well* (Student #9) and: *It was 2weeks classes for reading and writing and some skil which could help me in the exam* (Student #34) and a usual tactic of test takers, the past paper: *Self-study using the model question papers* (Student #40).

Preparation of English Language

Others however, took a more holistic approach to exam preparation and actually took steps to broaden their English language skills: *Reading classic literature and other worthwhile magazines such as Reader's Digest, National geographic, Daily newspapers and editorials; watching TV shows and movies to get used to the accent* (Student #4) and: *by reading a lot of English newspapers and watching English news and documentaries with sub-title on a regular basis* (Student #25).

Preparation for a specific discipline

Other students looked ahead to the post-test situation in their actual discipline and prepared accordingly: *I attended tutorials in my area of specialization* (Student #1) and: *I read online about the courses I was interested in and looked up the course material and borrowed books on similar subjects from the library* (Student #27) and: *Reading relevant materials relating to my course of study* (Student #45).

2. Experiencing the PTEA

Survey respondents were asked to what extent (on a Likert scale) they thought their PTEA scores matched their abilities in English in all the skill areas. The responses were: Very Well (n=13), Fairly Well (n=20), Not so Well (n=10), Not at All (n=3). Subsequent comments in the open questions that followed generated the following themes:

- Self-perception of proficiency not matching some PTEA scores
- Self-perception of proficiency matching some PTEA scores
- Computer medium of test affecting performance
- "Native user" proficiency
- Limited time as a factor in the test performance
- Inaccuracy of spelling score category

Self-perception of proficiency not matching some PTEA scores

Despite 33 out of 46 respondents claiming some satisfaction with their PTEA scores in the initial Likert-style question, there were more detailed comments and expressions of dissatisfaction in the open comments section: *In areas like reading, writing, oral fluency, pronunciation and listening. This area's do not match my ability in English language* (Student #10) and: *The reading and writing score sections are not a true representation of my abilities. I believe I should have gotten a higher grade than the grades I attained in these two sections. My spelling scores are also not a true representation of my spelling abilities. I am satisfied with the scores of the other sections* (Student #28).

Self-perception of proficiency matching some PTEA scores

Other respondents however were more accepting and satisfied with their PTEA scores: *I like to think that I am perfectly bilingual, and the test results seem to show it* (Student #37); *i think PTE was a reliable font of my English skills* (Student #18) and: *The test reflects and matches my ability very well in every area of the language* (Student #25).

Computer medium of the test affecting performance

One common theme in the survey was that of the medium of the test (computer) affecting the test performance of certain individuals: *In my view I can do better if the test was not on computer. This thing effects result* (Student #26). This type of response expressed the novelty of doing a test of this nature on a computer without human interaction: *Many people are doing this type of test of english for the first time (i mean the whole speaking to computer and*

stuff), so some people may even make some mistakes from lack of technical-knowhow. knowing how to use the headset, fast typing, some may have hearing impairment, etc (Student #45). It was expressed that the medium of the test was problematic for students from certain countries in that it: *might work against individuals who had had less exposure to computers: A paper-based test should also be provided for a class of test takers from developing countries where computer is not common for test validity and reliability [Student #7].* The concern was also raised that errors made operating the computer during the test could significantly affect test scores: *Student should be given chance to attend all questions, especially the unaswered questions, which he or she might have clicked or passed because of technical problem. For example, I mistakenly, cliked the next question and I lost more than 3 unaswered questions* (Student #41).

"Native user" proficiency

Some respondents revealed that they had a very good command of English from their education systems: *The education system in Kenya has a high standard of English language preparation which starts right from primary school to university level. This has made it easy for me to express myself and to understand both simple and complex English* (Student #28). Others who also identified with English still found the pace of the exam problematic: *I found the PTE exam's pace a little difficult for non-native English speakers. Although I have been educated in English right from primary school to graduation and even speak fluent English at home - even I agreed that the allotted time for some questions is too less* (Student #27). This sentiment links in with the next theme of time as a factor in test performance.

Limited Time as a factor in the test performance

Some respondents mentioned the limited time in which to take the test as a possible factor in test performance: *I had many surprises during the test, worst of which; if you do not respond in three seconds for speaking, you get moved to the next question and lose the score* (Student #13) and: *The listening section is especially difficult to keep up with!* (Student #27); *Reading - not enough time to read those loooong texts wich are king of hard* (Student #2); *the words were not on display long enough* (Student #46).

Inaccuracy of spelling score category

Questions were also raised as to the validity of some of the enabling score categories, in particular the spelling category: *spelling mistakes are eventually typing errors* (Student #20) and: *I am not however satisfied with the score on spelling because I was typing and the errors might have been 'keyboard' error in an attempt to complete the answer within the time given* (Student #7).

3. Use of the PTEA in Admissions

44 of the survey respondents indicated the PTEA or IELTS required entry score that had been outlined to them by their prospective universities (illustrated in table 62 below) and made further comments regarding the appropriateness of those scores (summarised in a thematic description below the table). According to the table, most people seemed happy with the entry score they had been asked to attain:

Table 62: Opinions on Appropriateness of PTEA Entry Score (or IELTS score*)

Scale of Opinion on Appropriateness of Entry Score Entry Score (GSE unless stated) + number of people indicating opinion and score (in parentheses) n=44				
Too low (1)	Low (3)	About right (29)	High (9)	Too high (2)
45 (1)	36 (1); 54 (1); 62 (1)	80 (1); 72 (1); 70 (2); 67 (1); 62 (1); 60 (4); 59 (3); 58 (1); 55 (2); 54 (1); 51 (2); 50 (3); 48 (1); 30 (1); *6.5 (3); *5.5 (1); B2 (1)	62 (1); 60 (1); 55 (2); 54 (1); 53 (1); 51 (1); 50 (1); 46 (1)	65 (1); 59 (1)

Impressions of Entry Scores Set "High/too High"

There were comments on the entry levels being set too high: *I am advising that there is need to review the PTE Academic pass mark so as to give room for many people willing to study in UK and elsewhere to open up there changes of getting admission* (Student #9 commenting on an entry score of PTEA 55). Moreover, some thought the high entry levels did not reflect the use of English in daily life: *I felt the entry score is a bit high because the test situation is not what someone encounters in everyday life (both academic or social) so a 55 entry score is a bit high* (Student #24). Others thought that the set scores were too high for speakers of other languages: *I believe the score is too high because not every region or country in the world use English on a regular basis. countries like Tanzania, Portugal, Angola, Congo just to mention but a few use other National languages other than English and expecting students from these countries to attain such a score is unfair* (Student #28 commenting on an entry score of PTEA 59). This was echoed in the comments of another student: *Because i believe there should be a kind of relief for those that are no English speaking countries* (Student #1 commenting on an entry score of PTEA 54)

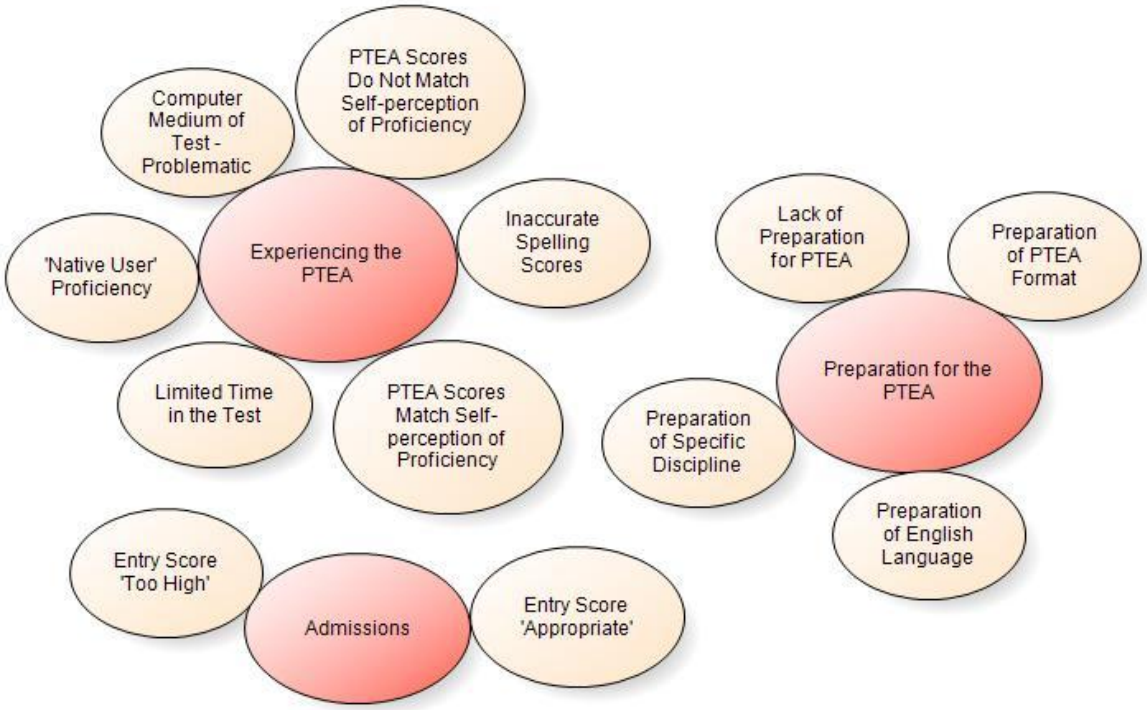
Impressions of appropriate entry score to course (about right)

However, a large number of respondents (n=29) were accepting of the entry score as appropriate: *For a postgraduate study, at least above 50 score is essential because one is to work independently* (Student #7 whose entry score was PTEA 55) and: *To understand the English speaking country, entry score is right because you have to communicate with the people of that country where you are going for your study* (Student #11 commenting on an entry score of PTEA 58). There was support for the entry threshold scores of below B2 (GSE 59)*: It is not very high score to have to show for an academic study, on my point of view is actually quite sensible. The school accepted my score and gave me a place, even though it was 54* (Student #39 commenting on an entry score of PTEA 55) and*: If a student can score above 50 in the Pearson English test, then he or she can cope with university degree programme taught in english* (Student #41 commenting on an entry score of PTEA 54) and: *Studying at university, it is essential that a student has a good grasp on the english language in order to succeed at their course. The PTE test score seems to reflect the level of english required in order to cope at university* (Student #15 commenting on an entry score of PTEA 60). For others, a "decent" PTEA entry score of 54 was considered low and meant that there would still be problems with language encountered in the UK*: this is because England is purly an english speaking country, so having an average Pte score means you still are stll to some extent deficient in one way or the other* (Student #45).

Only 22 out of 36 of the students who continued with the survey to the end said that they had actually entered university. Reasons for not getting in to university were due to the following factors: "visa problems" (n=2); "financial constraints" (n=3); "academic grades not good enough" (n=1); "PTEA scores not good enough" (n=4) and "admissions still in progress"

(n=3). The thematic map below illustrates the themes that emerged from the recruitment survey.

Figure 13: Summary of Recruitment Survey Themes



Overall the survey was successful in:

- Recruiting four participants for the case study
- Piloting the self-rating tool for language proficiency
- Gaining an impression of the perspectives of outer and expanding circle test takers, on their test preparation, test performance and proficiency.

4.2.2 The Case Study

4.2.2.1 Data Collection

Respondents agreeing to further tutorials were then taken on as the case study (n=4). The final data comprised the four cases' data from the recruitment survey, the three tutorials, and any samples of written work. The schedule for the case study was as follows: the recruitment survey took place from June to September/October 2013; the first tutorials were conducted between November and December 2013; the second tutorials were conducted between January and June 2014 and written assignments collected; the third tutorials were conducted and more assignments were collected between August and October 2014.

First Tutorials

Once respondents to the recruitment survey had indicated that they would like to be interviewed I set up the first tutorials. For this, I emailed the respondents from the recruitment survey who had left their contact details and arranged to meet them face-to-face or online using Skype technology. The tutorials consisted of semi-structured interviews (appendix 18) with each respondent (who now became a 'case') to:

- Gauge impressions of their PTEA performance
- Probe reasons for the self-rating choices made in the recruitment survey
- Use the "experiencing academia" cards to explore the cases' experiences of their first month or so at university and thus gather their perceptions of any linguistic or other difficulties encountered on their course so far (if any).

The Subsequent Tutorials

These followed the model above in using the "experiencing academia" cards within the tutorials to assess if any new problems had arisen since the last tutorial. An added focus to the second tutorial was to address writing as evidence of academic performance – to interview the cases about their written performance so far, and begin to collect some actual assignments from them for analysis. Thus, the second tutorial was preceded by an online "pre-tutorial questionnaire" (appendix 19), which had been sent to the cases via email and filled in online. This questionnaire provided an opportunity to prepare the cases for the second tutorial in terms of what I wanted to ask about written performance. In the questionnaires the cases were firstly invited to again rate their proficiency in the overall and four sub-skills (as they had in the recruitment survey) as part of the ongoing self-rating aspect of the study to see if their perceptions of their proficiency had changed. Then I adopted concepts on written performance from Banerjee (2003, p.427) to ask cases if they "never", "rarely", "sometimes", "often" or "very often" experienced difficulties in the various aspects of writing at university. Responses to these were followed up on at the start of the second tutorials. The survey also included a small section on self-assessing spoken performance such as in presentations. Cases were also asked to send details of the marking criteria from their department (how to interpret the marks of their assignments) as well as any individual feedback from tutors on their assignments.

This predictive validity study relies on evidence of performance. Thus, part of this evidence included written assignments of the cases (documentary evidence), to gauge how well linguistically they were "performing" in their subject and to be able to relate this performance back to their PTEA score profile of skills. Assessment of a cases's written performance took into consideration:

- The cases' original PTEA scores compared with self-rating of their skills
- The cases' coursework scores
- Identification of issues in writing from the online pre-second tutorial questionnaire
- Assessment of their written work from 3 sources (the university, first marker (myself) and a second marker (a colleague with EAP experience)

The third tutorial towards the end of their first year allowed me to ask the cases about their written work after having analysed some of their assignments, to get their reflections on what they felt they had achieved in the year, what they had learnt on their course and to address any outstanding issues or queries on their proficiency and performance as their first year came to an end. Some of the cases were finishing their Masters (one year) while others had another year of a diploma course to complete (leading to a Masters).

4.2.2.2 "Missing Data" – Tutors and Students (cases)

One data set that was unsuccessfully sought out was contributions from the cases' course tutors. Ideally, perspectives from course tutors are an invaluable extra dimension with which to assess cases' language proficiency and academic performance at university. However, gaining tutor's perspectives on individual students has proved problematic in previous studies (Banerjee, 2003) and it was suggested by Banerjee (ibid, p.126) that a more successful strategy might be to ask tutors about problems in general without reference to individual cases. I however did seek tutor opinions on the specific cases in my study, but unfortunately, I was simply unable to involve the tutors even after repeated attempts (contacted via the cases for ethical reasons). Reasons given were lack of time or tutors expressed no interest in taking part.

There were also important missing data sets from the case study. Case A proved quite elusive in that he did not participate in a second or third tutorial and did not provide written assignments despite repeated attempts to contact him by post, email, phone and various forms of social media. However, he did intermittently keep in touch and never formally withdrew his participation so there was always a chance (however slim) that he would respond at some point. Similarly, Case D did not provide examples of her written performance apart from an initial university diagnostic written assignment and she did not complete her third tutorial. Again, she did not actually formally withdraw her participation at any point. The cases were often simply too busy or preoccupied to return calls or emails. However, since Cases A and D did not express withdrawal from the research at any point, I have included data from their completed surveys and interviews in the study. Therefore, "incomplete" data was collected for these two individuals. The final degree transcripts and grades/degree awards were not collected from three of the cases as contact with participants was discontinued before this point in their studies apart from Case C.

4.2.2.3 Analytic Approach to Case Studies

I outlined the analytical approach to the tutorials (thematic analysis) in section 3.1.5 because the same process was applied to other data sets – the Pearson Country Representative Interviews and the Outer Circle Student Survey. Here I outline the analytical approach to the written assignments because it was unique to the case study. I used the grades and associated marking criteria/comments provided by the cases from their respective institutions and what they said about their writing experiences in the second and third tutorials to build a picture of their "written performance" at university. I compared the cases' comments to the institution feedback and in the analysis this was compared to their PTEA skills scores of writing and the

four relevant sub-skills of grammar, vocabulary, spelling and written discourse. In addition to these inputs, myself and a colleague commented on the writing of each assignment. This was done for each assignment that was received. I was "Marker 1" and my colleague was "Marker 2". My colleague has EAP tutoring experience at a university in London so he was familiar with assessing academic written work. We did not use any formal measure with which to assess the writing. Instead, I wanted to generate impressions of the linguistic proficiency in the writing assignments to compare with comments and reflections from the cases themselves (expressed in thematic descriptions of the interviews) which would be compared with institutional feedback and the PTEA Score profiles.

In instances where I had received assignments from the cases before the third tutorials, I introduced some "member-checking" referred to by Baxter and Jack (2008, p. 556) and Duff (2008:171). Member checking is a feature of longitudinal interviews and is usually the chance for the interviewee to comment on or respond to how the researcher interprets the words of an interviewee from a previous interview. This allows for ambiguity to be cleared up or themes and issues to be clarified. In my case, transcripts of interviews were not be sent to the cases, but I produced personalised third tutorial protocols according to what I had noticed in the written assignments and my assessment of them. This feedback was presented to the cases in the tutorials of students B and C and the cases responded to the feedback which is documented in the findings for each student (case) below.

Analysis of this data allowed me to determine if the case was struggling with any particular aspect of writing such as the structure, coherence, the grammar, or the content (vocabulary). Detailed comments on the assignments are tabulated in the findings for each case (tables 69 and 73 below).

4.2.2.4 Findings - Anonymous Participant Profiles and Actual Data Collected

The four case study participants (cases) comprised three Nigerian nationals and one Ghanaian none of whom indicated English as their L1. They were all studying in HE institutions in the UK, three of which were universities and one of which was a college which ran courses from which degrees were awarded (in association with universities). The cases' names and their HE institution names are not used in the description of the findings in the interests of ethics and confidentiality. Two of the cases were already on MSc courses and two were on Diplomas leading to MScs. Their profiles are summarised below in table 63. The first participant, Case A, provided self-reported PTEA scores only while Cases B, C and D all self-reported *and* provided an official transcript of their PTEA certificate. Their PTEA score profiles are displayed within the thematic descriptions below.

Table 63 Anonymous Participant Profiles (4 Cases)

	Case A	Case B	Case C	Case D
Gender	Male	Female	Male	Female
Age	33	25	24	46
Nationality/L1	Nigerian/Isoko	Ghanaian/Fante	Nigerian/Kalabari	Nigerian/Igbo
Course	Level 6 Graduate Diploma in Business Integrated Management (1st year), leading to an MSc. Fast-track-one-year)	MSc International Business (MIB)	Diploma in Petroleum Engineering leading to an MSc.	MSc in Human Resource Management & Training (Distance Learning)
Course Duration	March 2013-March 2015	Sep 2013 – Sep 2015	Feb 2013 – July 2014	Feb 2014 – Feb 2015

Table 64 summarises exactly what data was collected for each of the cases. As mentioned above, no tutors were recruited to comment on the cases' proficiency and academic performance, and Cases A and D did not do the full set of tutorials, nor provide written assignments for analysis.

Table 64: Summary of Data Collected for each Case A - D

Participant (Case)	Initial recruitment questionnaire including self-reported PTEA scores	Official PTEA Score Report seen?	1st Tutorial Interview	2nd Tutorial Interview	3rd Tutorial Interview	Examples of written work	Scores of assign ments & exams	Score criteria/inte rpretation from university
A	Yes	No	Yes	No	No	None obtained	No	No
B	Yes	Yes	Yes	Yes	Yes	2 assignments	Yes	Yes
C	Yes	Yes	Yes	Yes	Yes	3 assignments & 1 dissertation	Yes	Yes
D	Yes	Yes	Yes	Yes	No	1 Academic Writing Assignment (diagnostic)	No	No

I now present a summary of the biographical notes, score profiles, thematic descriptions (including direct quotes) and thematic maps of the tutorials (Braun & Clarke, 2006) for Cases A to D together with the presentation of the findings from analysis of their written assignments. Words in italics within the thematic descriptions denote the actual words used by the case and the number in brackets (T1, T2 or T3) denotes the tutorial in which the comments were made. The themes are presented under broader headings for each case and summarised in a thematic map at the end of each thematic description. I report the cases in order of the amount of data collected for each one. I start with Case B, then C, then D and then A.

The PTEA scores and CEFR levels are colour-coded to represent their alignment as follows:

Table 65: Colour coding of PTEA score ranges aligned to the CEFR

CEFR Level	C1+	B2	B1+	B1	A2	A1
PTE Academic Score Range	76+	59-75	51-58	43-58	30-42	10-29

Note: C1+ denotes a "C" level incorporating C2 but B1+ indicates a level in-between B1 and B2, according to the Pearson Score Guide (Pearson 2012a, p.40-41).

4.2.2.5 Findings - Case B

Case B was a 25-year old woman from Ghana. She was studying for an MSc in International Business from September 2013 to September 2015 at a school (HE institution) in the UK

affiliated with an institution in France. Her first language was stated as Fante. She had had work experience in English in Ghana in various roles (including the Ministry of Finance) for four years before coming to the UK.

Case B's PTEA score profile is below together with her self-rating scores.

Table 66: Case B - PTEA Scores and Self-rating of Proficiency

Overall and Communicative Skills Scores	PTEA Scores and CEFR	1st Self-Rating	2nd Self-Rating	3rd Self-Rating
Overall GSE	72	C1+	C1+	B2
Listening GSE	74	C1+	B2	B2
Reading GSE	73	C1+	B1+	B2
Speaking GSE	73	B2	B2	C1+
Writing GSE	73	C1+	C1+	B2
Enabling Skills Scores				
Grammar	63	n/a	n/a	n/a
Oral Fluency	73	n/a	n/a	n/a
Pronunciation	77	n/a	n/a	n/a
Spelling	41	n/a	n/a	n/a
Vocabulary	65	n/a	n/a	n/a
Written Discourse	90	n/a	n/a	n/a
Date of test/self-rating	23/05/2012	18/11/2013	01/04/2014	30/11/2014

Case B displays a score profile that is a consistent B2 (at the higher end of the band towards C1+) in all the communicative skills and almost all the enabling skill categories and she scores even higher on pronunciation at 77 GSE (C1) and the top score of 90 GSE in written discourse (C2) suggesting that Case B would not struggle at university as a result of a deficit in linguistic proficiency. The spelling score is a lot lower than the other skills but that score category along with the other enabling skills scores has a large error of measurement (Pearson, 2012a, p.43). Despite this score profile she was disappointed with her PTEA score and thought she could have done better in the test even though she selected "fairly well" when asked to what extent the scores were a reflection of her language ability. During her first self-rating she rated her proficiency to the descriptors aligning to the C1+ band of the

CEFR. There was a gap of about a year and a half between taking the PTEA and her first self-rating. She reported not feeling stressed doing the PTEA and felt her performance was much better than in a TOEFL-iBT test she had taken previously. Table 66 indicates that the level of proficiency indicated by the self-ratings decreased from the first, to the second, and then to the third tutorial by which time they matched more closely her PTEA score report. This could be to do with her encountering of issues in reading, writing and listening performance in her first year (see below).

Thematic Analysis

The main areas covered in the tutorials were: "Proficiency in English", "Experiencing the PTEA", and "Experiencing Academia".

Proficiency in English

Despite Case B's statement of Fante as an L1, it was clear through the themes such as "English in Ghana", "Previous Academic Experience" and "Previous Professional Experience" that English was her main language particularly in her educational and professional life: *I've been taught English since from nursery, so yeah we speak English in Ghana (T1)*. She had four years work experience in Ghana with English as the medium of communication including in the ministry of finance. For B, Ghana's relationship with English is an inheritance from her country's colonial past: *Ghana was colonised by The British…yeah so we've been speaking English for quite a long time (T2).* The use and appropriation of English in Ghana includes a rivalry in the region as to the quality of English: *Ghanaians and Nigerians are always arguing who speaks the best English (T2).* In terms of the self-rating of proficiency above, she had initially indicated a high regard for her own

proficiency as the descriptors she chose were at C1+ level but throughout the year it dropped until generally matching the PTEA scores (mostly at B2). As a comparison, she self-reported an 'A' grade in the WAEC (her CSELT) which is the top grade indicating a high proficiency in English (UCAS, 2014, p. 27) and this may have given her confidence in her initial self-rating in English.

Experiencing the PTEA

Taking into account her previous academic experiences, she showed some feelings of being affronted at having to take another English test, due to her high scores in English at school: *I think it's unfair, cos erm, during secondary school we write a lot of things in English and we learn English again, you know so, erm…if…if say I got an 'A' in secondary school I don't see why I should write the test again (T2).* However, she did acknowledge that those with lower grades may have to take additional international English language tests and reasons (perhaps erroneously) that economic migrants from her country overstaying might justify the use of language tests in admissions policies: *I think it is discrimination but I think we brought it on ourselves (T2).* In terms of test performance, she was disappointed in her score profile, which explained the high level given to herself in the first self-rating: *because I think I could have done better that's why I chose higher scores (T1).* She didn't understand why her spelling score was low, attributing the low score to typographical errors (in common with many comments in the recruitment survey) adding that she had been "in a rush" to complete the exercise and mentioned a lack of time as being a factor in her test performance (also chiming with a theme in the recruitment survey).

Experiencing Academia

According to Case B, language proficiency was not a factor in her academic performance: *I was very prepared in terms of language (T3),* but instead, issues to do with her performance revolved around gaining new knowledge and skills. This was represented in coding such as "presenting", "working with others", "assignment writing", "using I.T", "accessing online readings" and "adjusting to academic culture". Linguistically-speaking, Case B did have issues with reading speed in terms of encountering new academic vocabulary. Checking the meaning of some specialised vocabulary items had sometimes slowed her down but this was minimal and most of the reading difficulties stemmed from questions around accessing online resources: *in our school we have a lot of I.T. stuff… online books and online stuff and I find it very hard to look for stuff online (T1).* This was something she had had to adjust to as different from her time in Ghanaian academia during her first degree: *I've gotten used to it but I still prefer paper copies so that I can highlight some things and you know, cos I'm used to that in Ghana we just, everything is hard copy (T2).*

She found the amount of work tough to cope with at first but then seems to have managed in subsequent tutorials: *The workload hasn't changed but erm…they spread it out a lot for this semester, it's not as hectic as it was last semester (T2).* Her difficulties in preparing presentations were not due to a language deficit: *erm…in, in regards to the content and the structure of the words, no, quite sure (T1)* and were more to do with her personality such as shyness. Indeed, over the course of the three tutorials she gained confidence in the arena of speaking compared to her initial lack of confidence where her shyness or reluctance to present was evident: *just the thought of standing in front of people and talking made me nervous (T1)* and: *I don't ask a lot of questions in class - it's just me (T2).* This was, however,

being addressed by her gaining more experience in presenting: *it has improved my confidence and the way I speak has improved, yes (T2),* and by the third tutorial, the experience of working with a diverse student body in this regard was acknowledged: *because erm we had a lot of people from diverse backgrounds different countries and stuff, and we had a lot of group studies, and erm…getting to understand each other and work together to achieve a common goal was quite interesting, [Interviewer:: "ah"] yes so I think I learnt something from that as well yeah (T3).*

The sense of a gain in skills and knowledge of what was needed in academia was apparent in her identification of particular skills such as presenting, as above, and reading: *there's an art to reading a lot of material and picking the important points which I think er…I find quite hard (T2)* while other skills such as the conventions of referencing were acquired: *in terms of academic skills the referencing was quite hard and I, I learnt how to reference well here, in the UK (T3).* She also encountered new varieties of English among both staff and students which were a challenge for her: *the accents of the teachers, cos some of them are from I don't know like Indian, some different places so sometimes I find it hard to understand what they're trying to say (T2).* Overall, there was a sense that her language proficiency was sufficient but in terms of a skills gain she had benefited from the experiences on her course and in particular had self-reported an improvement in her reading skills, presentation skills and ability to work with others.

Dissatisfaction was the second prominent theme in her experience of academia. This was related to access to tutors and the institution itself in terms of its facilities and marketing to students: *I think they were just concentrating on selling the MA/BA and not really telling us the truth (T1).* Case B was often unimpressed with the tutors' teaching, accessibility and feedback practices: *they don't present their slides very well, and they don't go according to*

their slides, they just keep on jumping cos they have limited time, every lecturer has maybe fifteen hours or three hours so they are rushing to finish (T1); we are really deprived of a lot, even the lecturers are hard to find and talk to and stuff so you're on your own (T2). Even by the third tutorial she didn't have confidence in their feedback practices: *they just go through it, they don't really pay attention to any detail on your paper and then give you any grade they think (T3).* However, in terms of overall academic achievement, the disappointments in the institution faced during her studies were not determinants of her academic success. Her personality and ability to learn saw her succeed: *apart from being disappointed in the campus and everything I…I managed to you know, learn and get good grades, so I think I was focused on the whole, I didn't allow that to deter me from attaining my goal (T3).*

3. Academic Writing

Before the 2nd tutorial, which focused on writing and speaking experiences, in the pre-tutorial questionnaire (appendix 19), Case B indicated the following issues:

Table 67: Case B - self-selection of issues in academic writing and speaking

Degree of Difficulty	Areas Identified
Often has difficulty	Meeting the word limit
	Finding enough to write (content)
	Being understood clearly by my audience (sp.)
	Expressing my ideas confidently (sp.)
Sometimes has difficulty	Organising my ideas in an appropriate structure
	Expressing my ideas clearly
	Writing in an acceptable style
	Vocabulary (sp.)
Rarely has difficulty	Understanding the purpose of the assignment
	Writing a relevant answer
	Vocabulary problems
	Grammar Problems
	Spelling Problems
	Punctuation problems
	Using what I have read in my written work
	Pronunciation & Grammar (sp.)

Her coursework scores (marked out of 20 by her department) were as follows:

Table 68: Case B Written Assignments & Marks

Assignment Name	Mark (out of 20)
1. International Business	17 (best)
2. Corporate Finance	15.4
3. Intercultural and Ethical Issues in Business	15.3
4. Managing Information Systems	15.2
5. Research Methodology for International Managers	14.4
6. Marketing Management	13.5
7. International Marketing	13
8. International Negotiation	12.9
9. Global Trade Relations	12.5 (worst)

Two of these assignments (number 1 and number 9) were collected and the institutional mark and comments are accompanied by comments from Marker 1 and a Marker 2 (table 69).

Table 69: Case B - Written Assignments, Scores and Feedback

Assignment Name	Institution Mark/comments	Marker #1 Comments	Marker #2 Comments
Assignment #1 International Negotiation Assignment (worst)	12.90 - 'competent-good' 'The student demonstrates a fairly good grasp of the key concepts, knowledge, skills and abilities required for the module'	The only reason I can think that it was marked down is that content seems to have been limited Language proficiency itself does not appear to be a problem. The work reads well generally-speaking with minor errors. The mark (12.90) according to the criteria is one of her lowest but linguistic issues don't seem to be mentioned in the grade band 12-13 at which Case B was scored. I can only conclude that the content was the issue.	The writer obviously has a good command of English but is weaker when it comes to structuring an academic argument
Assignment #2 Management Information Systems (best)	15.2 – 'very good' "The student has a thorough knowledge of concepts and/or techniques with a fairly high degree of skill in the use of those concepts, techniques to satisfy the requirements of an assignment or course"	This is a very accurate linguistic piece of work. No discernable 'errors'. Again, I can only deduce that lost marks were to do with content and fulfilling the task criteria. Linguistic proficiency is not an issue according to this piece of work.	High level of linguistic range, accuracy and coherence… However, there is a lack of overall structure to the piece

Language proficiency issues were identified in the markers' comments in terms of the structure of the written work for the reader (marker 2) and insufficient content (marker 1). In terms of self-identification of writing issues, Case B also identified structure and content (finding enough to write) as areas for improvement in the pre-2nd tutorial questionnaire (table 67) and in some of her tutorial comments talked about what tutors had identified for feedback on her assignments: *the content and the style, and how it was structured, yeah (T1).* Her comments were in response to member checking during the third tutorial whereby I presented the 1st and 2nd marker comments to her. She conceded that content and structure were areas for improvement for her but she also said that she was not sure which areas exactly were lacking due to no personally written or face-to-face feedback from her tutors, which added to her sense of dissatisfaction with the tutoring on her course. Regarding issues with speaking (table 66 above), she attributed the problem of being understood when giving presentations to her accent and variety of English which may have been unfamiliar to her classmates: *maybe the intonation, yeah, mainly the intonation (T2).*

Progression through the three tutorials

In the first tutorial, her dissatisfaction with the reality of the institution was evident as well as the quality of her contact time with tutors. She expressed her shyness in presenting and her encounters with different varieties of English. A lot of the scenarios she highlighted as problematic were to do with academic skills. She did not express problems with language in any specific form. She expressed a willingness to access skills support in presenting if it was available in the college.

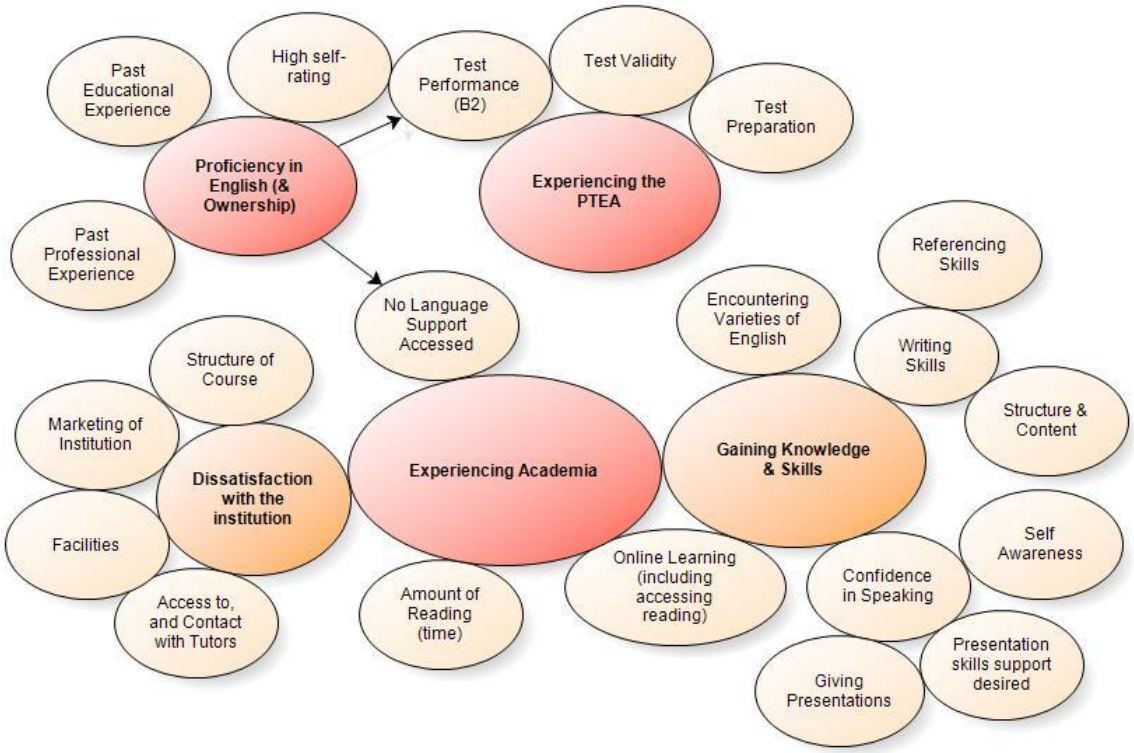
In the second interview, her issues were largely about her struggle with the workload, accessing readings and reading online, talking about improvements in her academic skills such as giving presentations and growth in confidence. She talked more about her written

work, marks, and why she thought she had received certain marks. She also expressed her apprehension towards the upcoming dissertation particularly in regards to content: *A bit apprehensive actually because I don't know what to write about to make up all those words (T2),* and she was also worried about adequate tutor support for the dissertation: *the problem is it depends cos he's coming in London in April so most of the work will be on email, and if she doesn't respond (agh), I'll be in trouble (laugh)… so even before I've started I'm already frustrated…yeah (T2).* She talked more about her use of English in Ghana, the politics of the test and how she had prepared for the PTEA. She expressed feelings of being affronted at doing the PTEA or of having to prove her English yet displayed an accepting attitude at having to do the test.

The third interview probed her views on her writing assignments and her overall feelings towards the course and institution (mainly her dissatisfaction with their facilities and student support) and her performance. She saw the bigger picture of the value of a Masters course, lamenting that she would have appreciated more time and opportunities to develop the wider skills to take away from a postgraduate course: *it's important because Masters is where you meet people, you network and meet people and learn a lot of stuff (T3)* which was an experience missing from her course. The overall impression from the tutorials and analysis of written work is that she already possessed the linguistic proficiency for university, which was reflected in and inferred from her PTEA scores. Her "academic performance" in terms of her written assignment marks was reasonably good throughout. What she gained on the degree course was an awareness of how she needed to perform on her course in terms of spoken tasks (presentations) and in improvements needed in her reading and written work and growth in confidence and the acquisition of specific academic and social skills such as referencing and working with others.

The thematic map below outlines the three main areas (in red). From the area of "Experiencing Academia" emerged two major themes (in orange) of "Dissatisfaction with the Institution" and "Gaining Knowledge & Skills". Her proficiency in English and experience in academia meant that she did not feel the need to access any language support at university.

Figure 14: Case B - Thematic Map of the Interviews



4.2.2.6 Findings - Case C

Case C was a 24-year old male from Nigeria studying a Diploma in Petroleum Engineering leading to an MSc over a two-year period from February 2013 to July 2014 at an institution in England. His first language was stated as Kalabari. Case C had a strong educational background in English including an undergraduate degree in the medium of English in Nigeria, but he had minimal work experience in comparison to the other three cases.

Case C's PTEA scores and self-rating choices are below.

Table 70: Case C - PTEA Scores and Self-rating of Proficiency

Overall and Communicative Skills Scores	**PTEA Scores and CEFR**	**1st Self-Rating**	**2nd Self-Rating**
Overall GSE	61	B1+	C1+
Listening GSE	64	C1+	B2
Reading GSE	53	B1	B1+
Speaking GSE	74	B1	B1+
Writing GSE	55	B1	B1
Enabling Skills Scores			
Grammar	45	n/a	n/a
Oral Fluency	76	n/a	n/a
Pronunciation	71	n/a	n/a
Spelling	56	n/a	n/a
Vocabulary	62	n/a	n/a
Written Discourse	40	n/a	n/a
Date of test/self-rating	23/11/2012	27/11/2013	24/03/2014

Case C has a score profile that suggests he might have had some issues with linguistic proficiency on his course of studies. His reading and writing scores (GSE 53 and 55 respectively) are at B1+ level just below B2 and just below Pearson's threshold of GSE 57 for a postgraduate course. As they are key skills for the researching and writing of assignments at university level, this suggests that he may have had some problems in these areas. His grammar (45 GSE) and written discourse (GSE 40) are significantly below B2 level. However he had a strong score profile in a number of categories and his overall score

of 61 was acceptable because his institution had set an entry score for the diploma leading to an MSc set at 54 GSE (B1+). His spoken and listening skills (including pronunciation and oral fluency) seem particularly strong in his profile. In his first self-rating he marked himself relatively low compared to his actual PTEA CEFR levels, particularly in speaking. In the second self-rating, he put a very high rating overall but lower in other areas. He self-rated his writing as low (B1) at both points which suggests some self-awareness of his shortcomings in that area. Even though he did do a third tutorial (via Skype), he did not fill in the self-rating survey that I sent to him. His reflections on self-rating are expressed in the next section. He was the only case whom I did not meet face-to-face. All his tutorials were conducted via Skype.

Thematic Analysis

The themes emanating from C's tutorials revolved around the three central areas of "Proficiency in English", his "Experiencing of the PTEA" and "Experiencing Academia", which concerned the main sub-themes of a gaining of skills and encountering issues in academic writing. These are described below.

Proficiency in English (& Ownership)

A certain proficiency in English can be seen in the educational background of Case C in the theme "previous academic experience". He had already studied in English at school in all his subjects and had acquired a Grade C6 in English language in the Nigerian WAEC qualification. A C6 is the 6th best grade in the WAEC system, one level up from a "pass" (UCAS, 2014, p. 27). He had pursued higher education in Nigeria in the medium of English, completing a BSc in Industrial Chemistry, so already had some background knowledge of his

subject – Petroleum Engineering. He had no professional experience in Nigeria after his BSc, but did do a two-month Petroleum industrial safety course. Regarding his L1, he stated it as Kalabari but in fact also used Nigerian pidgin English at home: *my parents speaking like, in a way you're speaking to me now in English, like speaking that you know, it's not necessary, they'll speak my local language, they'll speak pidgin English to me, it, it's very informal (T2).* Formal Nigerian English is officially used in public life including in universities: *my country, the first, my country first language is English (T1).* In terms of indications of ownership he indicated appropriation of English when comparing himself with expanding circle classmates: *I think you should interview someone whose, their first language isn't English, that would really be great for your… I have a friend from China – I will speak to him about it (T2).*

He seemed confident in his English in comparison to his classmates who were mainly from expanding circle countries: *I know for sure whether my English, my English is like, fairly better than most of them (T1).* Regarding his views on his country's qualifications as proof of proficiency, Case C did express confidence in his CSELT to indicate ability to communicate but did not express full confidence in Nigerian qualifications. Case C did question whether his own attained level in his CSELT had prepared him for study in the UK : *someone who's able to get a 'C' in WAEC definitely would not have any problem communicating with English…but coming over here to, to the United Kingdom you might need a…some more than that, some more than that (T3).* This statement in the last tutorial may have been a reflection on his own study experiences given his own score in the WAEC (C6). Despite this score, he did not report any problems communicating but did have problems in his written assignments (outlined below).

Experiencing the PTEA

He was quite disappointed in his test performance, even though he scored B2 in the test overall: *it's English and we speak it, we use it everyday so I...I wasn't expecting to do as badly as I did (T1)*. He was not satisfied with his reading score as it was the lowest of the four skills and he claimed he was better at reading than the test score may have suggested. However, in terms of his first self-rating, he actually rated himself *lower* than his PTEA scores in all of the categories except listening, with a gap of a year between taking the test and his first self-rating.

Overall, he did not feel that the scores reflected his ability although he did feel that the scores were acceptable for admission to his university course (set at 54 GSE). In terms of test preparation affecting his scores, he had had only one week to study the format and content of the test and as a result claimed to be unfamiliar with the content which could have led to him being under-prepared for the test: *Someone told me, someone told me that it wasn't going to be difficult and it was er…it was er…everyday, everyday, English (T2)*.

In common with some of the respondents in the recruitment survey, his attitude in preparing for the test could be interpreted as somewhat casual: *… I didn't see it as something to do…to to seriously prepare for (T2),* and he accredited his performance in the test to his unfamiliarity with the format and the electronic medium of the test: *I didn't know that you would type when the speaker comes, when, when they speak out I didn't, I didn't know you had to type fast (T1); I didn't know that that time had expired and another question… so I didn't know immediately it comes on, you just type, so sometimes maybe you don't know how*

the question maybe, maybe some people don't know how to use yeah the headphones, how to type fast, all the computer system so that also, can also effect someone's score (T1). These comments on timing and the actual taking of the test mirror those voiced by other test takers in the recruitment survey yet it is unclear how this actually affected his scores.

Experiencing Academia

Case C experienced academia in terms of the two main themes. He became self-aware of issues with his academic writing (including self-organisation and time management) and he gained lots of skills on his course especially in terms of group work in interacting with students from a wide variety of backgrounds. In terms of the "actual experience" scenarios presented to Case C in the tutorials, he indicated some difficulties in the following areas: "comfortable working in groups"; "ability to complete your written assignments"; "using IT."; "UK academic culture"; "language in textbooks and readings"; "understanding the language of instructors"; "writing for non-academic purposes"; "willing to ask for help from your lecturers or tutors".

Firstly, the issues in group work were concerned more with his interaction skills: *working in groups, there are, there are lots of things involved, you need, you need, you need to, be patient with some kind of people (T1)*. Included in this (as with Cases A and B) was an encountering and managing of new varieties of English among his expanding and outer circle classmates: *When someone who is not English is speaking to you, your level of concentration would, would change (T2)* and: *you can't prepare for that, just…you manage it, that's the right word (T1).* This encountering of varieties of English also applied to attuning to the English of the ('inner circle') UK*: the accents, the accent, the, the British accent...it's...they speak, they speak fast...and you have to really be keen, to have to listen keen (T1).* Issues in

working with tutors were more to do with the perceived status or power distance: *Er...you know definitely they are your lecturers so just...you won't be as free talk… communicating with them as you would be maybe communicating with maybe your mates...so…(T1*). It is unclear as to whether this was a cultural difference or a question of maturity as he was quite a young student. Certainly, his ability to work with tutors improved during the tutorials when by the third tutorial he was made class contact to represent the class, and by then he had improved his skills and confidence: *when you stay with, with people longer…when you stay with people longer period of time I think the interaction between you guys would become better (T2)* and towards his tutor: *I feel more comfortable approaching her (T2).*

However, his ability to complete written assignments still came down to poor time management and planning: *you see yourself still struggling, maybe two days, or even till the day of the deadline (T1)* and: *if the, if the assignment is 12 o'clock, I still see myself, still editing by 11:30, 11:45 (T2).* This had not been resolved by the time of the dissertation: *I actually had some problems during the final stage of doing the dissertation, I had some challenges putting my work together during the final stage of the dissertation (T3).* The challenges were that he had expected more input from his tutor and in his self-organisation: *two weeks to submission, two weeks to the deadline of the project when I was supposed to be really proof-reading and that was when I was able to get a hearing of how the project was supposed to go an…so I had to rush, rush, rush it (T3).*

In terms of his own proficiency in academia, in spoken English he displayed a confidence and a sense of ownership of the language in the way he talked about and compared himself to many of his classmates from countries such as Pakistan and Greece: *they were, were like 'ah, not English, no English, I can't speak English' (T2)*, which led him to conclude that his

English was better. This view of having better English was expressed in his language behaviour in academic situations - for example in using his initiative in group presentations which he had anticipated in the first tutorial: *if I am the one who will prepare the slides or if I'm among the group of people, who prepare the slides, doing the presentation wouldn't be a problem to me (T1).* By the second tutorial he had had some experience of group presentations which involved taking the initiative in order to get the task done: *I was the one who found the materials and distributed the materials, I just distributed, give everybody their parts to read and I even summarised, cos most of them didn't know what to do, so I summarised it and we sat together and made the, Powerpoint – this together (T2).*

Consequently, he had to cover for his group's mistakes: *so I was trying to protect the group and even when I came to speak I…tried to cover a few mistakes that other people did (T2).* However, he did see the need to improve his own English at times, which he dealt with by taking the initiative and responsibility through various behaviours including non-subject-specific conversations with classmates: *we have sometimes conversations to come up on politics and stuffs like that (T1);* being pro-active in class: *if I don't understand it, I ask questions in class (T1)*; and partaking in extra-curricular activities: *to improve my English, and English speaking and I joined a scrabble club…(T1)*

He had the communicative competence and personality with which to take part in these activities and to improve his skills. He used his initiative to read and write in his spare time beyond the academic demands of his course: *Yeah, reading for pleasure, relaxation, yes, I read, if you check my shelf you see a lot of books…all sorts of novels, yeah, and, and I'm also trying to write a novel…(T1)*. As with Case B, in relation to academic reading, a lack of vocabulary was a minor issue and subject-specific: *you have to stop reading, check the*

dictionary, know the meaning of the word before you continue reading so one or two times you encounter such words in textbooks (T1). He also recognised other skills that he had to develop on his course, in particular a need to be more autonomous and less passive, and less reliant on the tutors in a new academic culture: *during the project phase, the the guidance wasn't …the assistance given by the supervisor was not…perhaps like…the way it used to be back home the way it used to be back home when we have a project, you you, the feedback between you and your supervisor used to [inaudible] much more than this, they expect, I know it's masters degree but the, the, there's a, there's a, they expect you to do…they expect you to do… like, they expect you to work independent rather than getting ideas from your supervisor, they expect you to give ideas to your supervisor (T3).*

Academic Writing

Case C made the following selections in the pre-2nd tutorial online questionnaire:

Table 71: Case C self-selection of issues in academic writing & speaking

Degree of Difficulty	Areas Identified
Often Has difficulty	Writing a relevant answer
	Expressing my ideas clearly
	Writing in an acceptable style
	Meeting the word limit
	Finding enough to write (content)
	Grammar Problems
	Spelling Problems
	Punctuation problems
Sometimes has difficulty	Understanding the purpose of the assignment
	Organising my ideas in an appropriate structure
	Vocabulary problems
	Using what I have read in my written work
	Grammar, Vocabulary, Pronunciation (sp.)
Rarely has difficulty	Being understood clearly by my audience (sp.)
	Expressing my ideas confidently (sp.)

Selection of the issues where he "often" had difficulty (table 71) was quite revealing in that the type and number of issues he identified seemed to match what the PTEA had indicated in the score report, as the score categories "writing"; "grammar"; and "written discourse", were

the three lowest areas on his score profile. He also sometimes had difficulty with grammar, vocabulary and pronunciation in spoken performance but rarely had difficulties in other areas of speaking which reflected the high scores in those areas in the PTEA score report; for example, "oral fluency". Before looking at his comments on his written performance I also display his coursework scores and feedback on his assignments below. The marking guide of the institution was as follows: 70% =A; 60—69% = B; 50 – 59% = C; 49% or less = fail and their marking criteria can be seen in appendix 20. His assignment marks were as follows:

Table 72: Case C Assignments and Marks

Assignment Name	Mark (%)
1. Drilling Fluid	72% 'A'
2. Isomerisation Process	71% 'B'
3. Petroleum Hydrocarbons	66% 'B'
4. Deepwater Drilling	56% 'C'
5. Final Dissertation (Optimisation of Drilling Fluid Properties)	55% 'C'
6. Group presentation on a catalytic process in the oil and gas industry. (Steam Methane Reforming)	26/50 'C' (50% of the module)

Three of these assignments and his dissertation were collected and assessed as below:

Table 73: Case C - Written Assignments, Scores and Feedback

Assignment Name	Institution Mark/comments	Marker #1 Comments	Marker #2 Comments
Assignment #1: Drilling and Well Completion (Deepwater)	Total Marks: 56/100 Acceptable range of references. But referring to non-official resources (Wikipedia) Covering most issues related to deep water drilling, however some of them are not in appropriate depth. planning needs some	The writing is in the form of a report of a procedure and evaluation of that procedure. The layout and organisation seems clear. There is some lack of attention to detail (spelling). Use of passive is very important as it is a report and a report of a procedure, and he generally handles it very well with a few grammatical errors. Generally the meaning is not impeded.	High level of linguistic range, accuracy and coherence. A wide range of subject specific vocabulary used. Argument is strong in the section Problem of Deepwater Wells. The essay is very descriptive in the early sections. This might not be a problem in itself if description is what is required of this piece of writing. However, as a reader I am not sure if this is the case.

	improvement		
Assignment #2 Drilling and Well Completion (Drilling Fluid)	presentation needs to be improved Total Marks: 72/100 Few range of researched materials. Little Evidence use of scientific materials. Fairly good planning, however the paper ended without concluding remarks Fairly good understanding of the subject. Presentation of data needs is good, but needs some improvement. 71/100	It looks slightly better organised and presented than the previous assignment. He seems to have learnt from the feedback. From a linguistic point of view there seem few grammatical or lexical errors, those that occasionally crop up are minor and do not impede meaning or render the text nonsensical.	High level of linguistic range, accuracy and coherence. A wide range of subject specific vocabulary used. The essay sets out in detail different aspects of drilling fluids; it is a very descriptive piece of writing. However, from the outset there is no overall indication of the purpose of the essay or how the writer intends to structure the piece. Overall coherence is therefore not so strong.
Assignment #3: Isomerisation Process, Process Variables and Catalyst Selection	Good structure and presentation with fluent use of academic language	Again, a style for reports and describing processes. There are accurate passive structures used throughout. It seems a step-up from the first two assignments both in presentation and language. I would still say that his style is minimalist but perhaps this is indicative of the genre/subject	There is a high level of linguistic range and accuracy with a wide range of subject specific vocabulary. The writing is coherent and smooth flowing. Overall the text is very descriptive and relies heavily on secondary references. It is therefore difficult to see any clear argument developed. There is a lack of overall structure and coherence to the piece as there is no introduction, conclusion, discussion or signposting indicating how one section relates to the next – this had to be inferred from the essay title and table of contents.
Assignment #4 - Dissertation Optimisation of Drilling Fluid Properties	55/100 An acceptable and substantiated demonstration of understanding in all key areas of knowledge relevant to the work. Adequately structured and presented work, with clear use of academic language and reference to a sufficient range of relevant source materials.	Generally, a well-organised, structured piece of work with measured expressions and generally linguistically accurate sentences used throughout. There was no occasion where I seriously misunderstood because of linguistic reasons. What was still noticeable was a certain lack of elaboration especially in the discussion section, which was very short.	Clear and highly accurate; well-presented and structured. Ideas formulated well and discussed. Very well organised and coherent: easy to see where the writer is taking the reader and the overall purpose of the piece of writing. Clear and complex argument throughout. Consistent and highly accurate language used.

The scores seem reasonably good although he got a 56% for one assignment and 55% for his dissertation, which was quite disappointing for him. The second marker, however, rated his dissertation very highly in terms of organisation, content and language. The major areas for improvement were in the presentation and organisation of his written work. Regarding the structure and format of his writing in the tutorials he had initially mentioned issues with the

format of non-academic writing: *The format of writing, it's a problem, like letter writing, a formal letter…or writing, a, a, a, a C.V., or, or a proposal or something (T1).* He readily admitted to specific linguistic problems in his written English: *academic style, yeah, and grammatical errors (T1)* but he regarded any type of language support as unnecessary: *I didn't go, cos I saw it as a waste of time, because one, just, I just, I had lots of work so no need to be spending more hours in English so I didn't, I didn't go… I didn't need it, I didn't need it (T1).*

However, he did display an awareness of the intricacies of academic writing which he found "challenging": *you try to use your own language to, to give the same content of the materials you're reading from (T1)* and by the 2nd tutorial was aware of the experience he had gained in writing on his course: *I've written a lot of erm…exams and I've got an experience I've gotten some feedbacks, from lecturers and stuff and so, at least I'm, I have some…I have, experience (T2*). He saw his issues in writing as being to do with final editing, for which he found a specific solution: *I need to take it to someone who who can who understands English better than me to go through it so he can edit, edit the work (T1).* This suggested a lack of confidence in his ability to assess and edit his own assignments.

Feedback from his tutor on his assignments suggested that his planning and structure in assignments needed improvement, which he admitted to and indicated as having some difficulty with in table 73 above. For Case C, his personality and habits were a large factor in these shortcomings. For example, he confessed that he was at times "lazy" in writing and poorly organised in activities such as editing, proof-reading and planning assignments: *I was lazy about it, [Interviewer: "It was a bit rushed?"] I didn't plan, I didn't really plan about the assignment (T2).* However, this did not lead him to seek writing skills support: *I don't*

think I'll join any writing class or something, no (T1). By the second tutorial he had attended a compulsory "research and study skills" module but had not taken it seriously*: I didn't take it serious…I just…I, I attended the class but I didn't really take it seriously – it was last er… last semester yeah (T2) However,* by the third tutorial, he was able to look back and see what he had learnt about academic writing, a definite gain being the skill of referencing. In particular, the dissertation or final report was an eye-opener for him, and despite his low score on this he had learnt some lessons from the experience: *Yeah, now see, see the report, the writing style it was really poor probably like academic writing style it's during this project period that I actually learnt erm, academic writing how to, how to, erm reference properly and stuff like that and academic language is…during this report [inaudible] which was one key thing I learnt during the project (T3).*

Both the first and second markers made comments on the brevity and sometimes report-style of the assignments in terms of their minimalist content and style. When member checking with C in his third tutorial, he did say that this "report" style was expected in his discipline but admitted to his writing style being "really poor", saying he would do better if he had his time again, particularly regarding the adding of more references and sources to his work. His tutor had also given some indication in one assignment that C had at times been too brief and mentioned his "lack of concluding remarks" in one assignment. This all relates to "genre" and "meaning-making" in academic literacies and differences in expectations between tutors and students as to what constitutes writing in the discipline (discussed in chapter 5).

Overall, Case C demonstrated a proficiency and particular confidence in spoken English deriving from his previous academic and social background in his particular variety of English. This use of English was reflected in his confidence at university in the UK, especially in his interactions with classmates and his use of spoken English – something also

reflected in his PTEA score profile. He displayed a certain lack of skill with regard to preparing for and writing academic assignments, which he had become aware of during the course of the tutorials but he did display the language proficiency to develop strategies for dealing with problems encountered on his course. Regarding his experiences of university, his language proficiency did not seem to hinder his performance in his subject and he possessed a certain amount of confidence in his subject matter. His academic performance issues were largely to do with organising and editing his writing work and his poor time management. In terms of a lack of linguistic proficiency affecting his work he did admit to some difficulties with grammar, spelling, vocabulary and the content of what he was writing.

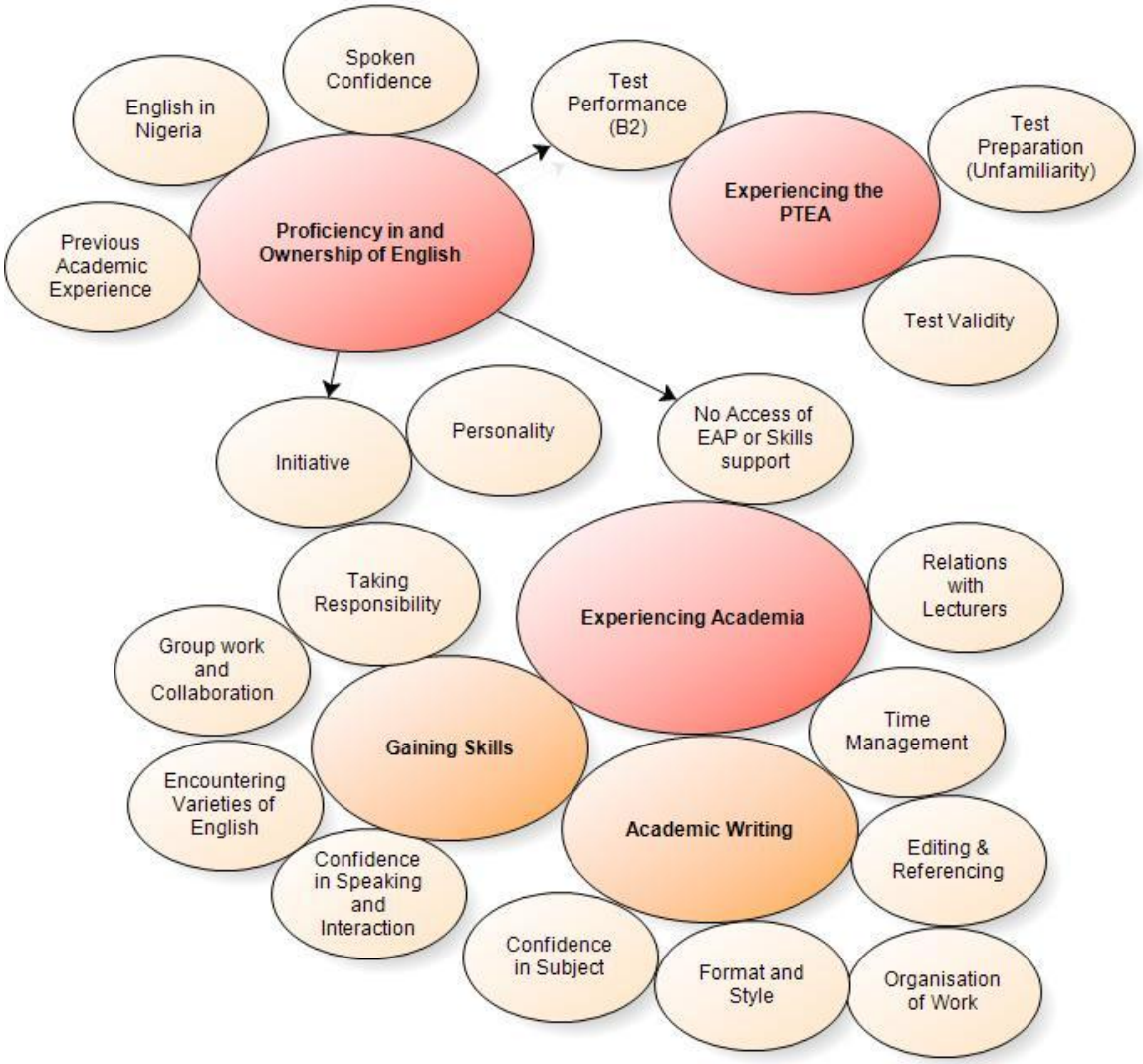
Regarding the changes observable though the three tutorials, Case C's self-rating scores indicate that he was aware of the deficiencies in his written work and performance, and the general trend through the three tutorials was an increasing awareness of his shortcomings in his academic writing and the development of improvement strategies. However, his written assignments continued to be affected by his relaxed and casual attitude in the actual planning, writing and editing of his assignments. His final 55% given for his dissertation is perhaps indicative of his continuing struggle with unresolved aspects of his proficiency such as that reflected in his low written discourse score in the PTEA enabling skills or his grappling with the academic literacy demands for his discipline. He sought feedback on his dissertation from tutors but I was unable to contact him further regarding this.

Regarding his views on the validity of a test such as PTEA as compared to the Nigerian CSELT (the WAEC) and previous academic experiences, by the third tutorial he reflected that students who had done an undergraduate degree in Nigeria in English (such as himself) would not need to demonstrate their English proficiency again but that Nigerian high school

students coming to do their undergraduate degrees abroad may need to take tests such as PTEA despite their qualifications in their national English tests (CSELTs): *Masters students who have already done a BA programme, I don't think, I don't think they need it… but maybe, people who are just finishing from high school and going to to er…first university may need it (T3).* These "caveats" are also mentioned by participants in the outer circle survey from strand 1.

A thematic map encapsulating the themes from Case C's tutorials is displayed below. His proficiency and confidence in spoken English in particular led to his taking the initiative in many instances at university and led to his gaining of skills as explained above. Where he was not performing sufficiently was in his writing and self-organisation, despite his confidence in his subject.

Figure 15: Case C - Thematic Map of the Interviews



4.2.2.7 Findings - Case D

Case D was a 46-year old female student from Nigeria studying a distance MSc in Human Resource Management & Training from February 2014 to February 2016 at a university in England (she was living in one city whilst doing her MSc at a university in another city, both within the UK). Her first language was stated as Igbo. She had just completed a 3-year BSc from a UK institution in Human Resource Management. She was a mature student, and before doing the BSc, had been working in business in Nigeria, which she continued be involved in whilst doing the distance MSc. Her UK BSc served as proof of English proficiency for the purpose of entry onto the MSc.

Case D's PTEA test result and self-rating of language proficiency is as below.

Table 74: Case D - PTEA Scores and Self-rating of Proficiency

Overall and Communicative Skills Scores	PTEA Scores and CEFR	1st Self-Rating	2nd Self-Rating
Overall GSE	45	B2	C1+
Listening GSE	52	B1+	C1+
Reading GSE	26	B1+	C1+
Speaking GSE	49	C1+	C1+
Writing GSE	50	B2	C1+
Enabling Skills Scores			
Grammar	47	n/a	n/a
Oral Fluency	51	n/a	n/a
Pronunciation	50	n/a	n/a
Spelling	61	n/a	n/a
Vocabulary	41	n/a	n/a
Written Discourse	22	n/a	n/a
Date of test/self-rating	16/12/2012	10/12/2013	30/06/2014

The score profile appears particularly weak with none of the scores, other than spelling, at the level of B2, despite her just having successfully completed a BSc in the UK. There is an obvious contrast with her first and second self-rating scores that put her proficiency much higher across the skills. There was a gap of a year between taking the PTEA and her first self-

rating, which may imply that there was a year of improvement in English skills to account for since taking the test. However, it transpired during the interview that she, along with all international students at her institution, had been obliged to take the PTEA in a hurry at the end of the last year of her BSc due to Tier 4 problems with the UKBA. This had been a stressful time for the students who were all completing final coursework. For Case D, the test performance on the day, possible unfamiliarity with the test format, coupled with the hurry in which she had to do the test could have accounted for her low scores in the PTEA. The conditions under which she sat the test may have affected her scores which leads to some doubt about her PTEA scores allowing me to infer anything about her actual language proficiency. By the second tutorial and self-rating (well into her MSc course) she seemed to have grown in confidence in her linguistic skills, rating them all at the C1 and above band.

Thematic Analysis

The themes from her two tutorials are discussed centring on her "Proficiency and Ownership in English", her "Experiences of the PTEA", "Experiencing Academia" and in particular "Gaining Confidence & Skills" and "Encountering Varieties of Englishes".

Proficiency in English (& Ownership)

Case D's proficiency in and ownership of English is rooted both in her educational experiences in the medium of English in Nigeria and the nature of English usage and status in multilingual Nigeria where English is the language of communication: *there are so many languages and there are different languages there, so that all of us understand English [Roy: "yeah"] so if you look at it you can probably say that English is our first language… probably that is why Nigeria erm…is very good in English (T1).* Case D received all her education in the medium of English both in Nigeria and the UK, had a diploma in secretarial

studies from Nigeria and experience in the arena of business in Nigeria where she used English in her then current role in her Human Resources company: *one I'm a mature student, secondly I did English in my WAEC and I did English again in er, information system management, I did English again in erm…my secretarial studies (T1)*. In terms of proficiency and ownership affecting the initial admissions into the UK, her CSELT (the WAEC), was sufficient to get her an interview for her undergraduate course: *they used my school certificate, because it's the equivalent of GCSEs in Nigeria (T1).* She self-reported a 'C' in her CSELT (the WAEC) but couldn't remember the exact C level (4, 5 or 6). Many UK universities generally see a 'C' level pass as equivalent to a SELT as indicated in the university admissions survey above. This acceptance of her CSELT was coupled together with an internal university English language proficiency test, which she passed. For her current postgraduate course her undergraduate degree was used as proof of language proficiency and not the PTEA.

Experiencing the PTEA

Case D's experience of the PTEA had a lot to do with the politics surrounding the use of SELTs in the UK. Her institution had been suddenly ordered by the UKBA to put its international students through the taking of a SELT in order to demonstrate their English proficiency. At the time she had been in the final year of her undergraduate degree in the UK. Her preparation for, and performance on the PTEA was thus generally a negative experience which did not perhaps reflect the proficiency and confidence she had gained on her undergraduate course: *To be quite honest, because of the pressure, and er, the problem, we are not concentrating because I was in my final year then…(T1)* In terms of her actual PTEA test performance, D found the actual environment of the test difficult: *with Pearson there's a lot of tension (T1),* in particular highlighting the booths in the test room and the fact of test-

takers being at different stages of the test during the exam. In relation to varieties of English, Case D stated a problem with the use of British accents in the test content implying she had not been used to those accents at the time. This was puzzling given that she had already been in the UK for at least three years at the time of taking the test: *sometimes for someone like I said who is coming in from erm…another country, to write an English test, they shouldn't expect that person to catch up immediately (T1).* However, in common with themes in the other surveys in this thesis and the other cases she also found the pace of the exam a problem: *I think erm, the speaking was okay, then the writing wasn't too okay, in short, to cram everything, it was because of that speed and accuracy there (T1).*

Case D claimed that her tutors and other British people would find the PTEA difficult and that being an international student was not the reason for poor performance on the test: *Let me be honest even the English they're giving us to write at that time even the British people too can fail it (T1).* Because of the circumstances surrounding her taking of the PTEA, it was difficult to ascertain whether her poor test scores were down to the stressful conditions in which she sat the test or whether it was a fair representation of her English proficiency at that time. As to why Nigerians took tests such as PTEA, unlike Cases A and B, she did not have any strong objections or feelings of being affronted, in fact she shared reasons why her compatriots might be obliged to take the test, reporting ethical problems with the Nigerian domestic exam system: *And probably again probably why they are doing it because our system is corrupt (T2)* and later: *it is not like here, sometimes a tutor can just come and put 'C', 'B' for you without meriting it (T2).*

Experiencing Academia

Because she was doing the distance MSc, she could not experience her postgraduate academic situation first-hand in the same way she had her undergraduate studies. Thus, in the tutorials she was able to relate back to a lot of her experiences on her recently completed undergraduate course rather than her MSc course, which she had yet to start formally at the time of the first tutorial. Gaining skills during her undergraduate course had been a big theme for Case D, in particular in her academic writing, group work and independent research skills. In working with others on her undergraduate course in the UK and in taking the PTEA, encountering different varieties of English had been a new and challenging experience. This difficulty in understanding other varieties of English applied to her difficulties with other outer circle African students (Zambian): *they are not fluent in the English (T1)* and expanding circle students such as Chinese: *they are very intelligent but they can't speak (T1).* This was a particular problem when it came to group work: *you find it so strange that sometimes when you find yourself into um, group presentation it's difficult to understand each other (T1).* As with Case C, this situation led to her having to take charge and show leadership: *my group members, probably they are not very good in English, I will have to do everything because it's a group mark (T1).* Regarding coping with local varieties of English in the UK (as with Case C) she also became aware of the differences between her own English and that of the host society: *The intonations and er, phonics is different from here (T1).*

A gain in skills on her undergraduate course was linked with her initial lack of confidence in the academic environment, perhaps from being a mature student who had been out of education for a while: *I didn't go to school for the past twenty something years before going back to school to do my degree programme so, and the method of teaching was totally*

different, the orientation and everything was totally different from Nigeria (T2). Adapting to the new academic culture had been difficult at first as she remembered how she had used to be: *sometimes I don't have confidence in myself, when they're communicating in the class, having an interaction, section so it was very very difficult I was like shying away and it was affecting my confidence, until I was challenged by one of my lecturers (T2).* As with cases B and C, she had also needed to acquire new skills in the academic environment in the UK, particularly independence in research, which she had adjusted to: *…then just go and then do the research yourself, I find it difficult a bit in my first year…But as time goes on then I keep catching up (T1).* Her confidence from her previous academic experience combined with her general language proficiency, personality and determination to succeed: *Okay I think all it takes determination, once you're determined to do it yeah, it has to do with determination, everything in life I guess it has to do with determination (T2).* She had gained a lot of confidence in her language and academic skills on her undergraduate degree and this prepared her well for her postgraduate course on which she now felt confident: *very confident yes (T2).* She had a real sense of having been transformed in the way she studied and thought: *even now when I went back home I couldn't fit in the life there, it's like I'm a different person (T2).*

Academic Writing

In the pre-2nd tutorial questionnaire on writing she had indicated the following difficulties:

Table 75: Case D Self-selected issue in academic writing (pre-2nd tutorial)

Degree of Difficulty	Areas Identified
Often has difficulty	NONE SELECTED
Sometimes has difficulty	Writing a relevant answer Writing in an acceptable style Vocabulary problems Grammar problems Punctuation problems

Rarely has difficulty	Understanding the purpose of the assignment Organising my ideas in an appropriate structure Expressing my ideas clearly Meeting the word limit Finding enough to write (content) Spelling problems Using what I have read in my written work

She had not experienced any speaking assignments on her distance course so those had not been relevant to her answers here (although she could have referred back to her BSc). However, as with Case C, in terms of academic writing Case D did seem to indicate some specific issues with language, in particular grammar: *You know sometimes we have to put in the prepositions, articles, [inaudible] sometimes, I struggle with it a bit (T2)* and with tense: *No, the grammar is like sometimes you know when I use past tense in the present tense, I confuse myself that's quickly I will realise…oh! That is not the right thing to use, I have to go back to the [inaudible], sometimes I struggle a bit (T2)*. This extended to specific uses of certain vocabulary items: *like for example we wanted to use like 'hence' 'as' how to use the break in a sentence (T2)* but again, like case C, she was quite pro-active and knew how to go about improving her own English: *it's something I identified myself even before the submission and because of that I, I have to buy the erm…'Common English" that's a book they call it 'Common English'…(T2).* These comments illustrate that the PTEA score report could reflect in some way her language proficiency, for instance in terms of her low scores for "Grammar" and "Vocabulary".

However, in writing assignments Case D generally claimed to have issues with structure and content rather than language use: *The feedback was okay erm most of the comments I got from them is er…it's the structure…the structures of the essay in most cases and more information (T1).* This was confirmed through observing the tutor feedback from the one assignment she had completed at the time of the tutorials – an academic writing skills (AWS)

task – an initial assignment particular to her institution. I did not gain access to the actual AWS task or any of her written assignments; however, I did receive the feedback she had received from the AWS task. The comments from her tutor (written) relating to writing skills and proficiency are as follows:

- "Your reflection on the use of Blackboard[11] was really clearly articulated"
- "A well-thought-out plan was presented in exercise 4 in order to address the proposed title"
- "Your attention to detail in completing exercises 5,7 and 8 is noted"
- "Within exercise 5 you demonstrate your eye for different writing styles and the difficulties with these varied styles"
- "The work submitted showed a gradual building of confidence with academic writing"

Areas for improvement included her referencing skills - a need to be more accurate in that regard and to be more evaluative of her readings as well as use of summarising and paraphrasing skills so as to develop more critical engagement skills. No specific comments were made regarding a lack of language proficiency *hindering* any aspect of her written work.

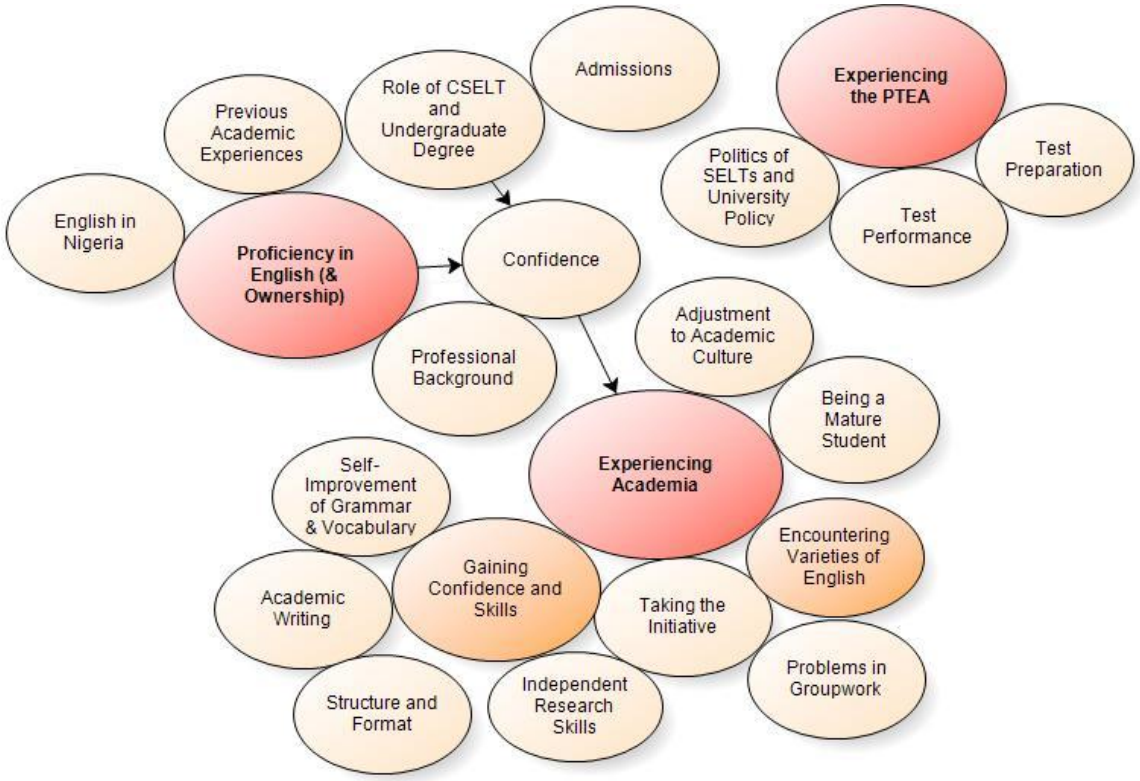
Generally, Case D possessed a lot of confidence from her experiences on her undergraduate degree course and her professional background as a practising businesswoman (to which her current qualification was directly relevant). Most of her adjustment to UK academic culture and language had seemingly been encountered during her experience as an undergraduate

[11] A virtual learning environment (VLE) and course management system

student. Despite having two tutorials, her distance MA was slow to get started and she had not, unfortunately completed any assignments by the second tutorial, apart from the AWS (above). The PTEA grammar and vocabulary skills scores in her score profile flagged up continuing issues in her proficiency, which were evident in her comments on academic writing in the two tutorials (above). Her self-rating scores (table 74) indicated that she felt quite confident in her linguistic preparedness for her MSc course. However, despite persistent efforts, it was impossible to make contact with her for a third tutorial when more detailed analysis of her written work and performance on her MSc course could have been conducted. Furthermore, longitudinal development of her progress in her course was not detectable because it was a distance course where written assignments would have been the main indicator of progress.

The thematic map below illustrates that Case D's confidence in English was fed by her previous proficiency in English from her professional life (business) and her educational background in English in Nigeria. Her undergraduate degree in the UK also contributed to this confidence. However she had also gained confidence in her academic writing and independent research skills on the BSc course she had just completed in the UK.

Figure 16: Case D - Thematic Map of the Tutorials



4.2.2.8 Findings - Case A

Case A was a 33-year old male from Nigeria. He was studying a Diploma/MSc in Business Integrated Management from spring 2013 to spring 2015. His first language was stated as Isoko. He had a business background and extensive use of English in previous professional contexts in Nigeria.

Case A's PTEA test result and self-rating of language proficiency is as below.

Table 76: Case A - PTEA Scores and Self-rating of Proficiency

Overall and Communicative Skills Scores	**PTEA Scores and CEFR**	**Self-Rating**
Overall GSE	71	C1
Listening GSE	71	C1
Reading GSE	69	B2
Speaking GSE	90	C1
Writing GSE	65	C1
Enabling Skills Scores		
Grammar	84	n/a
Oral Fluency	88	n/a
Pronunciation	90	n/a
Spelling	39	n/a
Vocabulary	51	n/a
Written Discourse	10	n/a
Date of test/self-rating	22/06/2012	18/11/2013

The score profile suggests that he is at a high level of English proficiency in all communicative skills, with a particular strength in spoken English. Possible weaknesses identified by the score profile seem to be in the enabling skills of spelling, vocabulary and written discourse. However, his overall written communicative skill score is high (65 GSE). His self-rating scores indicate that he has a high regard for his own proficiency. The main problem with Case A's score profile was that it was self-reported and no official transcript (though requested) was seen by the researcher.

Thematic Analysis

1. Proficiency in English (& Ownership)

Despite only being able to arrange one tutorial with Case A, proficiency in and ownership of English was illustrated by themes such as "English in Nigeria"; "Previous English Medium Education" and "Professional Experience in English". Case A comes from an educated urban African background where English is used as a lingua franca for communication in a highly multilingual society: *if you're speaking your own native language you need someone that…erm…shares similar language or values in order for you to communicate regularly, but everyday you communicate with English…* and: *the easiest way to communicate is still English.* The connection with the UK in terms of a historical relationship is put forward as a reason for not doing a test such as PTEA*: I would say…mmm…people under the, former, former…erm, British colony, former British colonies shouldn't really do this English test,* although he qualifies this sentiment admitting that there may be an urban-rural divide in terms of English proficiency: *although those in the cities of erm…these places actually speak good English but if you go maybe …down the line to the villages and the remote villages, erm yeah maybe then you can say they would do it* (i.e. would need to do a test such as the PTEA).

He had had an English-medium education in both primary and secondary school in Nigeria and experienced further education in the medium of English in Nigeria, with a Higher National Diploma[12] in Software Engineering: *Most of the things we learn are taught in English*. He had also had professional experience with marketing in Nigeria and worked for a

[12] A Higher Education qualification of the UK

firm that arranged study abroad for Nigerian students. He made a lot of statements reflecting his own comfort and ease with English which contributed to his "ownership" of the language despite his stated L1 : *I'm better in speaking in English than my own native language… it's easy to communicate, it's easy to understand….* Furthermore, his high self-rating scores, strong performance on the PTEA and non-access of language support at university suggested a confidence in his English ability. In the theme of "affronted", he expressed the view that individuals from Nigeria who have had all their schooling in English should not be asked to take tests such as these as it is seen as a form of repetition of their CSELT: *I don't feel they should go through the rigour of doing this test again*, a sentiment also expressed by Case B.

2. Experiencing the PTEA

When Case A was asked in the first tutorial about his poor scores in the areas highlighted above (written discourse and vocabulary), he made a number of comments that put this down to his performance on the test itself and in particular the time limitations in the test: *assuming erm we had erm some extra seconds, maybe twenty/thirty seconds extra, I'm sure I would have been able to do better in that…the timeframe, you have a limited timeframe.* These comments could also be reflective of the fact that he claimed to have had only 3 hours preparation for the test (in total), so familiarity with the test format could also have been a factor in his test performance. This compares to similar comments on timing and preparation for the test made not only by the other three cases above but also by others in the recruitment survey. However, whether or not he was familiar enough with the format of the test in order to perform well in it is not possible to infer from the score profile and his comments alone. In terms of self-rating, Case A rated himself higher than the PTEA scores suggested, at C1+ in every skill except Reading (B2).

3. Experiencing academia

As to his linguistic and academic performance, in the selection of "experiencing academia" from prompt cards within the tutorial, he had mentioned experiencing some difficulties in relation to the following areas: "UK academic culture"; "difficulty in completing the set work"; "finding resources"; "too great a quantity of materials to be studied"; "background knowledge/previous experience of the subject"; "understanding language of other students"; "speaking to other students". The latter two are encapsulated in the theme: "Encountering varieties of Englishes". He thought that UK accents were easy to understand but he sometimes found it hard to understand the Englishes of other international students. Although he declined to mention any specific variety, it was clear that he was encountering new accents and varieties of Englishes: *You know, some of the words they pronounce you're like…what?* Within the theme of "workload", he stated that the workload is heavier in the UK but that there are far fewer lectures to attend than in Nigeria.

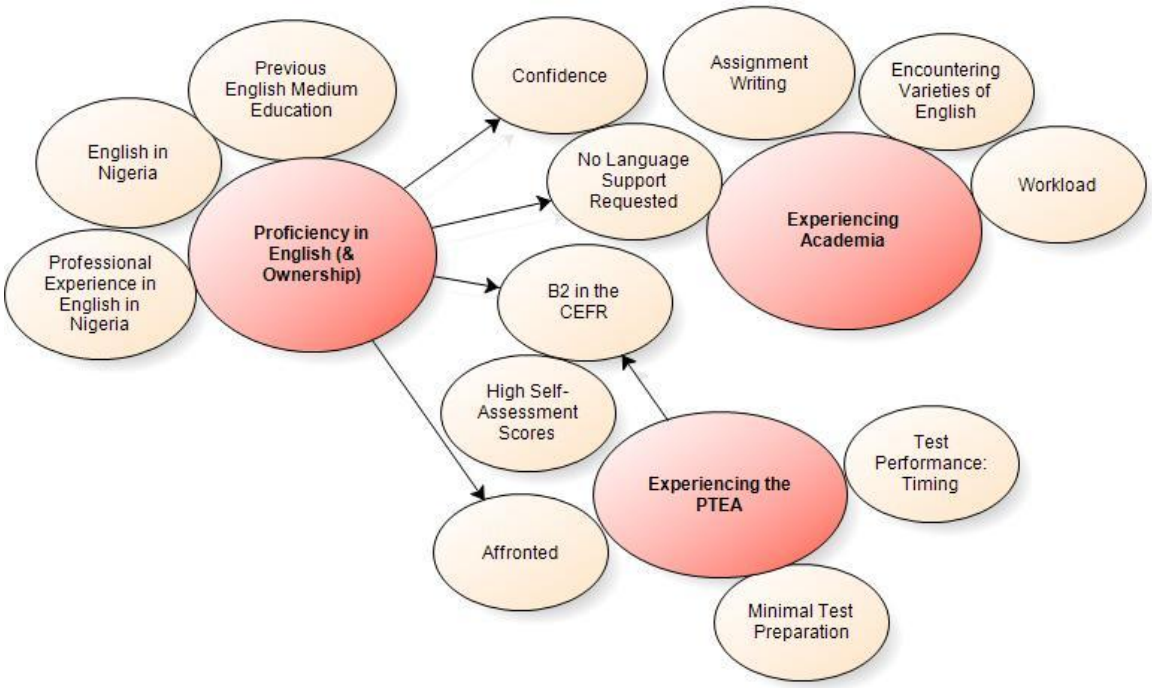
Regarding workload, he expressed the idea of wanting to pick and choose the topics within his set curriculum: *You know like…you look at the volume and like…am I supposed to? And, one of the other reasons is…um…in the real life…in the practical life, not everything we're learning you're going to put into the business practice, so only take what is necessary…and leave the rest.*

Regarding the actual assignments, he self-reported high scores for his written assignments: *I got a distinction…on one of the modules.* He commented on the feedback he got from his tutors on his written work where the feedback was not about his use of language but rather content issues: *It's just to do with content, not language…* The PTEA score profile suggested

that his written discourse and vocabulary may have been two areas in which he needed to focus when looking at his written work; however, he did not participate any further in the research and so it was impossible for the researcher to independently analyse his written assignments or obtain feedback and marks from tutors. Consequently, it was not possible to make any definitive comments on how his linguistic proficiency affected his academic performance.

The tutorial is summarised by the thematic map below. The three central areas of investigation in the tutorials are highlighted in red; proficiency in and ownership of English; experiencing the PTEA and experiencing academia. These are surrounded by the sub-themes described in the thematic description above. Case A's proficiency in English is linked with the fact that he did not request language support in academia (as was true of all the other cases). He expressed confidence in communicating in English and overall scored at the higher end of the CEFR 'B2' level. He also expressed objections at being asked to take the PTEA (expressed in the theme "affronted") and this is connected to his feelings of ownership of English expressed through his proficiency and background in English.

Figure 17: Case A - Thematic Map of the Tutorial



Chapter 5: Discussion

5.1 Introduction

From the findings in this study, two types of table for comparing findings and themes were drawn up for my own reference in preparing the discussion. Cross-case tables and a joint display of data illustrated which themes and findings were comparable and contradictory across the cases and other data sets. These informed the discussion which now follows. I first discuss the findings in relation to each research question (restated below), in particular I interpret them in relation to the literature and discuss the implications of the data. After this, I then go on to discuss the two main strands of the study in terms of: a) how the findings can be interpreted as regards the English proficiency and ownership of English of outer circle students and; b) how the findings can be interpreted as regards the predictive validity of the PTEA for the cases in my study. Both these strands address the wider focus of assessing the validity of outer circle Englishes for entrance to universities in the inner circle.

5.2 Answering the Research Questions

5.2.1 Strand One

5.2.1.1 Research Question 1: ADMISSIONS

What do the admissions policies of universities regarding nationality and CSELTs qualifications (as proof of English proficiency) suggest about the ownership of English and English language proficiency of outer circle students?

The data around this question shed light on whether country-specific English language tests (CSELTs) such as the West African Examination Council (WAEC) qualifications were accepted as proof of proficiency in English by UK universities. Despite the UKVI policy to

recommend (but not impose) SELTs, the acceptance of CSELTs by many UK universities in my findings (over 57%) indicates that many UK universities recognise that many students from the outer circle have already experienced education in the medium of English and have acquired a certain corresponding level of proficiency. Swansea University is an illustration of this orientation to CSELTs: *you will not be required to take a secure English language test if you have been taught in English at secondary level or above in one of the following countries and have obtained the qualification and minimum grade listed* (Swansea, 2015). Whilst this is a clear recognition of prior and sufficient proficiency in English, it is also contrasted with the fact that 54 of the 132 institutions surveyed (41%) did not accept CSELTs as alternatives to SELTs. As regards other data sets, this divide was also seen in the outer circle student survey whereby five of the 11 participants who were at university in the UK had had their CSELTs accepted as proof of English, and five out of 11 had been asked to do a SELT despite their CSELT qualifications.

This polarised response to CSELTs could be interpreted to mean that some UK universities either do not trust a CSELT as a valid indicator of English proficiency or that they do not recognise the prior proof of proficiency of outer circle test takers. However, non-acceptance of CSELTs may be due to other reasons such as reluctance to displease the UKVI in the light of what has happened to other institutions (Havergal, 2015; Shepherd, 2013). In those cases the UKVI found that the institutions concerned had admitted students with insufficient English or "invalid or "questionable" qualifications. The universities' refusal of CSELTs could also be a desire to simplify the admissions process regarding proof of English proficiency. Case D certainly experienced the ambiguity and confusion regarding her hurried taking of the PTEA at the end of her last year of her BSc course.

Specific reasons for rejecting CSELTs were not explored in the document analysis as reasons were not given in print in the online documents. However, from analysis of other data sets, there were concerns raised by certain individuals as to the reliability of CSELT qualifications (in particular the Pearson Nigerian and Malaysian representatives and some individuals in the outer circle survey and cases C and D). The Pearson representatives pointed to the more *comprehensive* measure of oral and aural communicative skills tested for in the PTEA, which could be currently lacking in some CSELTs. However, the WAEC was considered by the Nigerian representative to be a very effective measure of writing and reading proficiency – which are key competencies for academia.

Many individuals in the outer circle survey with CSELT qualifications and experiences of education in the medium of English expressed objections when these were undervalued by university admissions. On the other hand, Case C queried whether a 'C' pass in the WAEC would be enough for university entrance and Case D questioned the quality control surrounding the awarding of marks in the WAEC which relates to the issue of "questionable" qualifications reported by Havergal (2015), although in the case referred to by Havergal doubt was raised about an IAET. However, I would interpret the findings of the UK universities admissions survey as indication of a wide acceptance of a variety of CSELTs among UK universities. This in turn is recognition of outer circle proficiency in English and an indirect acknowledgement of the historical inheritance of English in many outer circle countries manifested in their current English language proficiency qualifications.

5.2.1.2 Research Question 2: PROFICIENCY DIFFERENCE

i) Is there a difference in the proficiency of outer circle test takers compared to expanding circle test takers as reported by PTEA and other academic test scores (IELTS and TOEFL-iBT)?

ii) Is there a difference in proficiency, as reported by PTEA test scores, according to the L1 of the test taker? (English L1 vs Other L1)

The descriptive and inferential statistics did indicate that as a group, the outer circle countries displayed statistically significant higher levels of proficiency than the expanding circle group (table 26). The large-scale test data suggested that there are significant differences in language proficiency between the outer and expanding circle as a whole but that when looking at individual countries and regions the picture of proficiency painted by the test scores is varied both within and outside the various circles. Some expanding circle countries display a high proficiency as reported by test scores, in many cases higher than that of certain outer circle countries. However, Singapore and South Africa display particularly high proficiency scores on IAETs, and outer circle African test takers also display consistently high mean scores at the required level for university entrance set by the UKVI (B2) which is statistically significant in comparison to both the expanding circle and outer circle Asia.

There were, however, significant differences in outer circle proficiency, with particular relevance for my case study when looking at Nigeria. Despite scoring high in mean scores for writing at 61.38 GSE (B2), Nigerians had a mean score of 55.48 GSE overall (B1+), lower than other OC African countries' mean scores (table 24) and just below the Pearson recommended score for entry onto postgraduate courses (GSE 57). Within OC Africa,

proficiency in English is slightly lower in Nigerian PTEA test takers as reported by mean test scores as compared to other outer circle African countries. Ghanaian test takers (table 37) with or without English as an L1 perform well (with mean scores of B2) on the PTEA (table 37) and similar findings were found for East Africans, South Africans and other Southern Africans such as Zimbabweans who also demonstrated mean scores at the requisite level for academia without it being dependent on L1 (tables 36, 40 and 41 respectively). These findings indicate that many African test takers are theoretically already at a level of proficiency in English suitable for university level study.

Outer circle Asia was more complex regarding proficiency levels. For OC Asia, as seen in the descriptive statistics of the sample, OC Asian languages were much more prominent than English as L1s because OC Asian countries are more likely to have an Asian L1 functioning as a lingua franca or majority language (Cantonese in Hong Kong; Bengali in Bangladesh; a variety of languages in India; Malay, Tamil and a variety of Chinese languages in Malaysia; Urdu, Pashto, Punjabi and others in Pakistan and Tamil or Sinhalese in Sri Lanka). This in itself is a reflection of the vitality of these languages in the social life of these nations, for example see the "Cantonese vitality" phenomenon in Hong Kong (Pierson, 1998, p. 91) which explains the lack of Hong Kong test takers putting English as an L1 - too few with which to perform a t-test on a comparison of mean test scores.

Regarding the expanding circle, many individual expanding circle countries match or exceed the proficiency in English of many outer circle countries as reported by PTEA, IELTS and TOEFL-iBT (tables 23, 50, 55 & 56). Thus, high mean scores aligned to B2 and above in the CEFR are not only applicable to outer circle countries, for example, South American and European countries display high mean scores in both the TOEFL-iBT and IELTS 2013 data.

However, as a group, the expanding circle had statistically significantly lower mean test scores and there is a core group of countries from the expanding circle Middle East and East Asia such as Saudi Arabia, Iran, China and Japan whose PTEA mean scores indicate a much lower proficiency level than the outer circle group means across the score categories (table 23).

The nature of English proficiency is therefore differentiated between the outer circle and expanding circle and also within the outer circle as referred to in the literature (Bruthiaux, 2003; Jenkins, 2014). Despite this variation within and between the circles, there are still important implications for the outer circle. Based on the Pearson claims about B2 and "coping" with the language demands at university and after analysing the academic experiences of my cases below, the statistical data suggests that students from countries such as Singapore, South Africa, other African countries, The Philippines, India and Malaysia all display at least a B2 level of proficiency with statistically significant mean scores which can lead to the inference that they would not be disadvantaged in academia due to a language deficit.

The TOEFL-iBT (tables 53 and 54) also suggested this is the case and African and Asian countries score at the high end of the proficiency scale within their continental groups in comparison to expanding circle African and Asian countries. This suggests that they could be thought of as ESL rather than EFL in terms of proficiency as seen in other studies (Zheng & Wei 2014, Bridgeman et al, 2015) where "outer circle" Indian PTEA test takers were seen to outperform "expanding circle" Chinese PTEA test takers and mean scores were significantly higher in every score category. Singapore and South Africa in particular have very high mean scores in the TOEFL-iBT test data for 2013 (tables 53 and 54). Together with the PTEA

score data, this further explains the concerns of test takers from those two countries regarding having to do IAETs despite their proficiency in English (UKCISA, 2011, p.16-17).

When L1 and test performance was looked into, evidence from the descriptive statistics suggested that the proportion of test takers in many outer circle countries and regions stating English as their L1 was very high, including: Singaporeans (88.8%); East Africans (81.6%); South Africans (80.7%); Zimbabweans (77.8%); and Nigerians (70.2%) (table 28). Stating English as an L1 is a self-identification of English as a first language, a clear demonstration of appropriation or ownership. The fact that Singapore has the highest percentage of L1 English speakers in its sample is supported in the literature, where it is stated that English is a de facto L1 for many Singaporeans despite official government policy (Lim, 2015; Tan, 2014). For outer circle Africa, there is evidence of a language shift to English happening in society (De Klerk & Bosch, 1998; Igboanusi & Wolf, 2009; Mesthrie, 1992; Offiong & Mensah, 2012) and an association between social class, education and proficiency in English within Africa (Kamwangamalu & Tovares, 2016; Matiki, 2001; Michieka, 2005; Ochieng, 2015) which means that many applicants to universities in inner circle countries will come from this social group.

However, when looking at Nigerian test performance according to L1 (table 38) it is apparent that those with English as an L1 do not necessarily perform better overall on the PTEA than speakers of other Nigerian L1s such as either Igbo or Hausa, whereas all three L1 groups perform well in writing (all above 60 on the GSE at B2 in the CEFR). The implication is that Nigerian test takers of the PTEA (as a group) are not necessarily at the B2 threshold overall for university study in English in terms of an overall score and therefore may not necessarily demonstrate a sufficient level of proficiency for university simply by being an “outer circle”

country. Their lower mean scores as a country imply that students from Nigeria may suffer a slight deficit in terms of linguistic "catch up" at university in the inner circle. This is significant for the cases in my study, three of whom were Nigerian and two of whom (Cases C and D) did express some difficulties with academic writing that were both of a linguistic and non-linguistic nature (see below).

As regards L1 and test performance in OC Asia there were varied results. Indians, Malaysians and Filipinos with English as L1s did score in the B2 band of the CEFR (tables 43, 44 and 46 respectively) although, on the GSE, Malaysians with L1 English performed higher on the PTEA (around 70 on the GSE in each category) than Indians and Filipinos with L1 English (around 60 on the GSE in each category). Within Malaysia, Malaysians with English as an L1 perform significantly better on the PTEA than those Malaysians with either Malay or Chinese Languages as L1s (table 44), indicating that Malaysians with English as an L1 would be less likely to struggle in academia due to language issues than the other L1 groups. However, there are indications for other OC Asian test taking nations that English as an L1 is not a significant factor in high test performance as Indians with L1s such as Hindi and Tamil and Filipinos with Tagalog as an L1 also scored highly on the PTEA - overall at just about B2 level in the CEFR (tables 43 & 46). Conversely, neither was English as an L1 a guarantor of high proficiency in other OC Asian countries (Bangladesh, Pakistan and Sri Lanka in tables 42, 45 and 47). Proficiency in English in OC Asia presents, therefore, a complex picture with contradictory information emerging from mean test score data according to nationality and L1s from Asian countries. However, within Asia, Singapore does emerge as highly proficient in English with regards mean test scores, with the vast majority of test takers in the Singaporean sample stating English as their L1 and the country displaying high mean scores on the PTEA and TOEFL-iBT tests at C1+ level on the CEFR.

The implication here is that outer circle Asia is far from uniform when it comes to the legacy of English and the current situation regarding proficiency. In some parts of outer circle Asia for example, English proficiency is high from the colonial inheritance such as in Goa (India), Singapore and among certain segments of the population in Malaysia (Bolton & Ng, 2014; Kurzon, 2004; Rubdy et al., 2008; Tan, 2014) whereas English can be regarded as more of an elite language in places such as India and Hong Kong (D'Souza, 2001; Pierson, 1998).

Overall, from the sample used in this study, outer circle African countries do have higher proportions of L1 English speakers compared to outer circle Asian countries (bar Singapore) within their test-taking population and in comparison to expanding circle countries. In terms of variation in mean scores according to L1, for the entire data set, having English as an L1 was statistically significant in having higher mean test scores in all score categories than those with other L1s but the same pattern can be seen in both the separate EC data and OC data. In OC Africa however, test takers seemed to do well on the test (B2) regardless of their L1 which suggests a proficiency in English across L1 groups and could be a reflection of the multilingual nature of Africa where many people have a significant proficiency in English despite another stated L1. There could also be a problem in the data here in that it was not possible for test takers to state more than one L1, that is, they could not state if they were bilingual and so the conclusions that I can draw from the data as regards L1 may be limited because the concept of L1 may be problematic and inaccurate as it is difficult to ascertain what the test takers may have interpreted as L1. For example, identification of an L1 can often be an emotive issue and test takers may be considering identity instead of proficiency when they identify their L1s.

Whilst acknowledging these limitations, when looking at mean test scores, nationality and L1 appear to be significant as test taker characteristics. The PTEA seems to differentiate the proficiency of test takers according to nationality and L1. For researchers such as Wang et al (2012), this difference in test performance according to nationality and L1 may be problematic. Wang et al refer to the difference in mean test scores on the PTEA of Indian and Chinese test takers in Zheng and Wei's study (2013) and state that this finding "calls into question the appropriateness of using the same set of item parameters for the two groups and the meaning of their score interpretations" (2012, p.612). If Wang et al are suggesting that the Indian test takers should not be doing the same proficiency test as the Chinese test takers then this would have interesting implications for those setting the tests - for example would tests such as the PTEA only be used for EFL learners of English? Would there have to be a separate test for ESL learners and further still, one even for "native speakers" in order to assess their academic English as opposed to none as a presumption of their proficiency in a general English? I argue that this differentiation in test performance according to nationality and L1 is actually confirmation of differences in the history, use and function of English in the societies of different groups such as outer circle Africa, outer circle Asia and the expanding circle (Gupta, 2006; Higgins, 2003; Kachru, 1985; Obaidul Hamid, 2014; Widdowson, 1994). In this way, the test is simply doing its job by revealing these differences in proficiency.

As a caveat, the somewhat contradictory data on test performance on the IELTS and TOEFL-iBT tests does throw some doubt on the ability of mean test scores to explain proficiency differences between outer and expanding circle test takers in studies such as this. For example, according to my data, some expanding circle countries perform 'better' in tests such as IELTS and TOEFL-iBT than they do on the PTEA (in so far as the tests are aligned

with the CEFR). This alignment could be the reason for this variation. Another factor contributing to a variation in test performance according to which exam is taken is the concept of self-selecting bias occurring. For example, students at a low proficiency from some nationalities may choose one test over another with which to test their English because of perceptions of one test being easier than another. Thus, test performance according to nationality is not necessarily a commentary on the proficiency of all test takers from a certain country. Further research would need to look at who is taking the test in each country and why. For example, my mean test score data for Ghanaian test takers from 2014 contrasts with findings on Ghanaian test takers of the PTEA in a study of Ghanaian school teachers (Stoffelsma & De Jong, 2015) where Ghanaian mean test scores for PTEA reading were reported at 46.6 (B1) for those taking the test between 2010 and 2012 (no sample size given) whereas my Ghanaian sample from 2014 (n=73) recorded a much higher reading average of 61.02 (B2). This all points to more comprehensive data sets needed on test taker background, nationality, L1 and test performance from large-scale academic tests such as TOEFL-iBT, IELTS and PTE Academic. It would also be revealing to process this data for all test takers of particular nationalities from when the PTEA began in 2009. This highlights that there is a need to investigate who individuals are, and why they have taken the test because not everyone who takes a test of this nature is doing it to study abroad, some will be taking it to assess their level of English, some might be taking it inappropriately (they are not the target test taker), and only some may be taking it for study abroad.

Another factor revealed by the recruitment survey is the medium of the test and test conditions (figure 13). The survey revealed divergent opinions and views, some respondents were content with their scores while others did not recognise their proficiency in the scores. This may be related to the phenomenon of "inflated self-assessment" (Kruger & Dunning,

2009) but it may also be a legitimate concern. For example, regarding the test conditions, some respondents cited the computer as a potential problem and difficulties with the format of the test (in particular the time limitations during the exam) but it was not possible to make a definitive connection between these issues and their eventual scores. However, more research in how these factors do affect test performance would be an important direction in future research and test takers may have legitimate concerns regarding these factors.

The results also support studies theorising a high and sufficient language proficiency among outer circle students at university, for example among Kenyan students (Gichura, 2010), which is supported by the wider literature on outer circle proficiency (Davies, 2013, Higgins, 2003). However, a higher proficiency does not necessarily equate to a higher academic performance in terms of final academic grades as some of those studies also revealed. Other factors impinging on "academic success" or performance are covered in the discussion of the case study data in the second strand and include a student's academic socialisation into their institution, and induction into the academic literacies of their discipline.

5.2.1.3 Research Question 3: PERCEPTIONS OF PROFICIENCY-OWNERSHIP

'What are stakeholder perceptions of outer circle proficiency and ownership of English for university entrance?'

The views of Pearson country representatives and outer circle students produced themes that complemented and contradicted each other and matched themes in other data sets. From the comparison of themes from both data sets I discuss these findings under the four headings below: the nature of outer circle proficiency; ownership of English; contradictory views

towards CSELTs as indicators of proficiency; and test preparation and format affecting scores:

The nature of outer circle proficiency

What was apparent from analysis of the Pearson interviews and outer circle survey data was that access to English-medium schooling favours wealthier middle class families and may be significant in determining L1 English usage. All the respondents in the outer circle survey were well-educated with prior qualifications in English and proof of proficiency on CSELTs. This was also matched in the educational profiles of the cases in the case study. Pearson country representatives from all three countries talked about the English-medium international schools and the ability of the middle classes to access these schools for their children. This is a trend in countries such as South Africa, Malaysia, Ghana, Singapore, Kenya, Nigeria, Zimbabwe and India with large middle classes. These same countries displayed high numbers of L1 English speakers in their test taking population (table 27) and displayed high proficiency in English in the form of PTEA scores in my study (tables 35-46). Proficiency is often based on educational level and therefore often "middle-class" because educational experiences in the medium of English in many outer circle countries are dependent on wealth and privilege (Kamwangamalu & Tovares, 2016; Matiki, 2001; Michieka, 2005; Ochieng, 2015).

Moreover, if a test taker from the outer circle does not have English as an L1 from birth there is evidence to suggest that they acquire it through education as indicated by Davies (2011a, 2013) who proposes that the "native user" can acquire or exceed the proficiency of a "native speaker" through education - the concept of the "idealised native speaker". The Pearson

Indian representative also alluded to this aspect of acquisition of English behind the ambitious drive of many test takers in India to pass the test for purposes of migrating abroad to work. In this way, English in India is seen very much as a language to get ahead in education and in work as in much of the outer and expanding circles. This is also seen in the case of the Malaysian representative interview where it was mentioned that many middle class Malaysians send their children to English-medium schools.

Ownership of English

Proficiency in English in the outer circle is bound up with aspects of ownership in the outer circle for example the idea of the "legitimate speaker" of a language (Higgins, 2003, p.620, Norton, 1997). From the Pearson interview and outer circle survey data this ownership was often expressed in the form of objections to taking tests. Many individuals from various countries in the outer circle survey from both outer circle Africa and Asia, expressed objections at having to take IAETs as proof of their English proficiency - again reflecting the concerns in the UKCISA survey (2011, p.16-17) which indicates an ownership of English. The Nigerian Pearson representative expressed the idea that Nigerian parents and teachers would voice objections to students having to do extra English proficiency tests, an idea also expressed by the Malaysian representative in relation to students.

Ownership of English was also reflected in comments about trends in society and touched on the issue of language shift. In terms of English as an L1 there was an acknowledgement from the Malaysian Pearson representative that those with English as an L1 may comprise as much as 10% of the Malaysian population – and that this is increasing (which compares with 33.3% of the 500 Malaysian test takers in the PTEA 2014 data with English as an L1). The Nigerian

representative mentioned the fact that English is an L1 in many Nigerian homes as did Case A. Marriage trends such as couples from two different L1s creates situations whereby the children are often brought up in English and this pattern was mentioned in particular by some cases from Cameroon in the outer circle survey. This "language shift" occurring in Africa is also reflected in the literature (Igboanusi & Wolf, 2009; Offiong & Mensah, 2012).

Contradictory Views Towards CSELTs as Indicators of Proficiency

As mentioned above in the discussion on the university admissions survey (RQ1 – section 5.2.1.1), CSELTs are not always seen as comprehensive tests of communicative ability in terms of covering the four skills of reading, writing, listening, and speaking. The Nigerian and Malaysian representatives both indicated that CSELTs are not as comprehensive as the PTEA in terms of measuring oral and aural communicative skills in English but in the Nigerian case, the WAEC is thought of as measuring reading and writing skills very well. The implication was that the PTEA is a more "valid" test from the point of view of the Pearson representatives, but the Nigerian representative did not state this categorically and in fact highly praised the written component of the WAEC. Some individuals in the OC survey and cases felt that a high school qualification in English (a CSELT) would be sufficient proof of English proficiency for entry to UK universities whilst others argued for an undergraduate degree in English as the minimum. This difference in opinion on the suitability of CSELTs as measurements of proficiency for university study links with similar doubts raised in the case study by Cases C and D. Doubts have been raised in the literature as to the reliability of outer circle pedagogy in English (Howell et al., 2012, p. 33) particularly with regards the "form of local English" in countries such as India and Nigeria. This may lead some to conclude that IAETS are necessary as a type of "quality control" for qualifications whose validity is in

doubt. However, more empirical research would need to be done on the comparability of CSELTs to IAETs in this regard.

Test Preparation and Format Affecting Scores

A uniform theme from the Pearson representatives was that familiarity with the format of the test was important in performing well on the test (also reflected in the concerns raised in the recruitment survey). From analysis of the recruitment survey and case study in strand two, many test takers were surprised or concerned about: the pace of the test; the test conditions; medium of the test (computer); and aspects such as not being able to go back to check previous answers on the test. The participants in the recruitment survey in strand two admitted to a minimal amount of preparation, particularly in terms of awareness of the format of the test, and to the encountering of difficulties with the content and test conditions - despite such information widely-available from Pearson on how to prepare (Pearson, 2012d). A lack of familiarity with the format and medium of the test could mean that the test scores would not reflect the actual proficiency of the test taker. Regarding the computer medium of the test, the Indian and Nigerian representatives voiced concerns about the familiarity of test takers from more rural areas with this technology. This concern with the test format and medium can be investigated in another study as it was not part of my study's investigation but it is reflected in the concerns of individuals in other studies such as the Indian test takers in Zheng and Wei's study (2014, p.138).

In terms of improvement in proficiency as preparation for the test, the Pearson Nigerian and Indian representatives mentioned that students' preparation for the PTEA was generally only in terms of familiarisation with the format of the test. They suggested that Nigerian and

Indian test takers do not see themselves as needing to improve their language proficiency as preparation for the test as they already see themselves as having become proficient from their schooling in the medium of English. The implication was that the test is seen as an "extra" measure of their proficiency for the bureaucracy of studying abroad.

The interpretation of the findings from the Pearson representative interviews and outer circle survey is that there is a high outer circle proficiency in English often based upon social class and educational experience in the medium of English. Ownership of English is intertwined with this proficiency (Higgins, 2003; Norton, 1997) but it may lie somewhere along a continuum or a degree of ownership such as that expressed by Seilhamer (2015, p.375).

The implication of this identification of sufficient proficiency in English and ownership of English among many outer circle students is that the validity of the test for outer circle students could be called into question. The statistical tests and thematic analysis point to an already acquired "native user" status among outer circle test takers. Whether an IAET such as PTEA or IELTS is suitable for many individuals from outer circle countries is a question touching on issues of test fairness (Kunnan, 2000, 2010). Significant "personal characteristics" such as nationality and L1 can have "systematic effects" on test scores (Bachman, 2004, p.156) as seen in the statistically significant t-test and ANOVA in this study. If Singaporeans, South Africans and other outer circle Africans are consistently scoring so highly on the PTEA and TOEFL-iBT, this may lead some to question whether the test is suitable for them. On one hand, the test may be a valid recognition of their high linguistic proficiency, but on the other there may be legitimate ethical questions raised as to why individuals such as these are taking these tests (UKCISA, 2011). The solution could be for universities to continue to choose their own method of determining the proficiency of

outer circle students without the use of IAETs and to introduce their own tests which could provide more useful assessment of a student's academic readiness for the skills and academic literacies needed at university. In other words, without knowing who the test takers are and why they are doing the test, it is not known whether the test may be being used for its intended audience or purpose.

5.2.2 Strand Two

In this strand I discuss and interpret the predictive validity of the PTEA in relation to how the students experienced academia and performed in academia. I do this, not only in relation to whether the test scores can be interpreted or related to any "deficit" in their linguistic proficiency or academic "study skills", but also whether this had any effect on other aspects of performance such as their engagement with academic literacy practices on their courses and I also interpret whether other variables had any effect on their academic performance.

5.2.2.1 Research Question 4: PREDICTIVE VALIDITY

How do the PTEA score profiles of individual students compare to their actual language behaviour and academic performance in their first year?

Language Proficiency

Regarding language proficiency and academic performance, this question was concerned with whether the PTEA score profiles allowed me to infer whether the cases were linguistically prepared for academia and whether the test scores indicated areas where the students did encounter a "cost" (Banerjee, 2003). This idea of a "cost" can also be seen in the evidence that without a solid general proficiency in English, students will be left feeling

"peripheral, isolated and unfulfilled" (Murray, 2010, p.58) for the other challenges of proficiency in academic literacy.

In the Pearson claims at the start of this study, a general proficiency at level B2 is stated as sufficient for university study (Minimum of GSE 59) and B1 insufficient. There is also a claim that a minimum of GSE 57 is the minimum threshold for postgraduate study and that the test is for non-native speakers of English.

In terms of how the four cases' score profiles measured up against these claims, Case B was only 'deficient' in spelling (GSE 41); Case A was potentially deficient in spelling (GSE 39), vocabulary (GSE 51) and written discourse (GSE 10), while Case C was flagged up in reading (GSE 53), writing (GSE 55), grammar (GSE 45), spelling (GSE 56) and written discourse (GSE 40) and Case D in all areas bar spelling (GSE 61). However, in the case of D, it should be borne in mind that she did report an unusual test-taking performance under difficult circumstances so questions arise as to how reliable the score profile is as a reflection of her actual proficiency.

With regard to the low spelling scores in three of the four cases' score profiles, it is difficult to infer anything from these as dissatisfaction with this score category was a common theme from the recruitment survey on experiences of the PTEA. Furthermore, Pearson themselves mentioned in personal communication with the researcher (De Jong, John, 01/06/2014), that some of the Enabling Score categories might be less reliable due to a potentially large error of measurement and this is also indicated in the literature (Pearson 2012a, p.43). However, the enabling skills scores can still be used for inferring possible deficiencies in language proficiency.

In terms of encountering problems to do with language which could be related back to the PTEA score profile, the cases did indeed encounter some language issues on their courses. In academic writing, Participants C and D were quite open about experiencing structural, grammatical or vocabulary problems. These problem areas matched their lower PTEA scores in the score categories of "grammar", "vocabulary" and "written discourse" (enabling skills). In the case of C and D, their low written discourse skill was also reflected in the feedback on structure and coherence (organisation) mentioned by the markers and the students themselves. As seen in appendix 6, the written discourse score is defined by Pearson as:

> Correct and communicatively efficient production of written language at the textual level. Written discourse skills are represented in the structure of a written text, its internal coherence, logical development and the range of linguistic resources used to express meaning precisely (Pearson, 2012a, p. 3)

The feedback on the assignment marks from their institutions, added to by the comments from the two markers of the written work, and the comments of cases B, C and D all identified the structuring of a written text and coherence as areas for improvement for the three individuals concerned. These issues of text organisation or "structure" as a factor in academic writing are mirrored in the findings of previous studies (Brown, 2008, Woodrow, 2006, Banerjee, 2003, Hill et al, 1999, Ferguson & White, 1994). In the way that the institutions themselves marked the assignments, language proficiency on its own was not specifically mentioned as part of their marking criteria though there are general phrases referring to proficiency within their criteria with descriptors such as "fluent use of academic language" (appendix 20). Difficulties the cases had in writing could also have been down to differences in expectations between tutors and students as to "what is involved in student writing" (Lea & Street, p.159) in the idea of "contested meaning".

In reference to the above, in adjusting to the demands of academia, there is a strong sense in which the cases generally had issues to do with academic writing that are also experienced by "native" speaker students and concerned referencing, editing, time management, workload, content, style, essay structure, and sourcing and using texts for essay writing. However, these issues were not identified as being caused by a lack of linguistic proficiency. "Native" speaker students could equally have made comments of the kind made by the cases with regards to their academic writing experiences and indeed problems of this nature in adjusting to academia have been documented for inner circle students (Lowe & Cook, 2003; Byrne & Flood, 2005). The implication here is that academic literacy is a concern for students from any L1 background including English.

In terms of self-awareness of linguistic or skills deficiencies, I found that the cases were quite aware of what they needed to improve upon. Banerjee's 2003 study suggested that her participants would sometimes cite cultural problems as a foil or "scapegoat" for what were really linguistic problems (ibid, p.389) but this was not apparent in my study. In talking about their proficiency for university my cases reflected findings from Ingram & Bayliss (2007) whereby participants were aware of their linguistic proficiency:

> The fact that participant self-assessment of linguistic proficiency closely matched actual IELTS scores suggests that students are more often than not aware of their proficiency and how it relates to their ability to perform different tasks. (Ingram & Bayliss, 2007, p.59)

From analysing the data in the tutorials, cases B, C and D in my study showed great awareness of their shortcomings and strengths in written English and in particular academic skills such as independent research or in their ability to gather and process materials for reading in preparation for writing assignments. Their self-rating of language ability may not always have matched their PTEA scores but their comments in the interviews did indicate a

strong awareness of their strengths and weaknesses in academic skills. A lack of matching between the cases' self-rating of proficiency to PTEA scores (tables 65, 69, 73 & 75) is perhaps indicative of the phenomenon in self-assessment whereby it is common for individuals to "over-estimate" their ability (Kruger & Dunning, 2009) but it could also indicate a certain self-confidence, legitimate belief in, and knowledge about, their own proficiency.

In terms of other conceptualisations of proficiency, Case C's eventual engagement and comfort with talking to and interacting with his tutor is an example of Murray's "professional communication skills" (2010, p.59-60) as is Case B's eventual confidence in presentation skills, and Case B and C's experiences in group work and working with individuals from other cultures and L1s. The implication here is that these students all had the requisite general English proficiency - which not only allowed them to function on their courses but also allowed them to develop and grow in their confidence and proficiency in these other skills. Skills of this nature are also identified as needed for home students (Byrne & Flood, 2005; Lowe & Cook, 2003).

Academic Literacies

For all the cases, adjusting to academia in the UK included acquisition of new skills which depend on reaching a proficiency at a threshold of general English, but also cross over into academic literacies: getting to grips with the academic conventions of writing in a subject; reading for writing (how to process readings in preparing and composing your own written assignments); developing independent research skills; adapting to new technologies such as accessing electronic copies of readings; adopting conventions and practices such as

referencing; and being proactive in managing workload and time. Case B for example mentioned the difficulty in adjusting to all the readings being in "pdf" form or online rather than the physical library she was used to in Ghana. This aspect of adjustment to academia is also mentioned by Brown & Holloway regarding Russian and African students adapting to accessing "Learning Resources" in a new academic environment (2008, p.12). Cases C and D noted how they had had to be more mature or proactive in their relationship with their tutors and better in self-organisation, factors also mentioned in the findings of Fox (2004); Kerstjen & Nery (2000); Lowe & Cook (2003) and Hill et al (1999). A student's ability to work independently on their own initiative is a key skill, but one which is an important skill for all students to acquire, both home and international.

Cases B and C mention practices which align with the notion of academic literacies. For example, this can be seen in Case B's comments on there being an "art" to reading materials and "picking the important points", and her comments on the "bigger picture" of being in academia - "masters is where you meet people, you network…and learn a lot of stuff".
In terms of reading, there is a sense in which a lack of language proficiency could be a reason for citing reading as an issue, for example a lack of proficiency would slow a student down considerably but this was not detected as a reason for a lack of academic progress among the four cases.

Indeed, students' problems with structuring and coherence in their written work can also been seen in the context of "meaning-making" and contested meanings of what it means to write in academia and for a particular genre "the linguistic features of structure and argument are clearly open to interpretation, and what may indicate argument for one person (e.g. cohesive ties, juxtaposition, reference, connectives) may not appear so to another" (Lea & Street, 1998,

p.167). Case B did not know why she had received low marks for some of her assignments and this can be related to the feedback practices of the university and findings from previous studies whereby students had thought they had submitted well-structured and coherent work but this clashed with the feedback given by the university and hence the students "often felt confused and unsure of what they had done wrong" (Lea and Street, 1998, p.164).

This may also explain any differences or discrepancies in the comments from the institution and the two markers in this study (tables 69 and 73) because individual tutors may also carry with them contested meanings and differing expectations of what makes writing in academia and in a particular discipline (Lea & Street, 1998, p.163). Furthermore, while the institutional tutor is the one familiar with the needs and expectations of the discipline and presumably assessed assignments according to set criteria (appendix 20) the other two markers commented on general language proficiency in terms of "atomised" aspects or "surface features" such as grammar or spelling as well as features of proficiency in the assignments such as cohesion, style, content and student voice. In Banerjee's study (2003) she mentioned that in feedback on student's work, tutors "rarely mentioned linguistic issues in isolation" so it is no surprise if the institutional tutors' feedback on Case B and C's work did not focus on linguistic issues. Adjusting to the demands of a particular discipline as well as a British academic environment is an experience that all international students must go through (Banerjee, 2003, p.388) and this adaptation or adjustment may not be related to the nationality or L1 of the student (Lowe & Cook, 2003, Brown & Holloway 2008). However, Banerjee cites language proficiency as the key variable in a student's ability to be flexible regarding adjusting to a new academic culture (2003, p.389).

My study of the proficiency and ownership of English of four outer circle students indicates that for these four individuals, already schooled and socialised in the medium of English, there was no indication of a serious deficit of language proficiency as a factor in their academic performance at university. This was a result of comparing their actual experiences at university to their PTEA score profiles relating to a B2 competency in the CEFR. However, the students all had to adopt study skills as part of their academic socialisation into the academic community, and to develop literacy practices such as meeting expectations for the institution in terms of producing and constructing suitable pieces of writing.

The implication of this is that it can be stated that the PTEA is useful for assessing the "atomised skills" of proficiency and for the cases it seemed to predict some areas of "deficit" or "cost" in the students' score profiles which matched with "deficits" identified in the proficiency of Cases B and C. However, what the test scores did not do was to allow inferences to be made about how these cases would negotiate and acquire the academic literacy demanded of them in their disciplines. The findings also show that my cases were all largely "proficient" students yet they had issues with study skills and academic literacy which affect inner circle students as well. This suggests that proficiency in terms of being academically literate is something for all students at university to work on and develop, not just international students who are seen as having a "deficit" in their language proficiency (Wingate & Tribble, 2012). None of the cases accessed EAP classes – they either did not have time or did not think these classes as useful to them. It may also have been the case that these classes did not offer them any solutions to their issues with academic literacy as articulated above.

5.2.2.2 Research Question 5: VARIABLES

What are the multiple variables that affect the academic performance of these particular outer circle test takers on their chosen courses?

The case study also revealed the existence of other factors such as affective factors and acquisition of skills that were important for their academic performance or success which are addressed here. However, there is some cross-over with the previous section.

Studies such as those by Kerstjens & Nery, (2000); Cotton & Conrow (1998); Hill et al (1999); Fox, 2004); and many others, suggested that there are multiple factors aside from sufficient language proficiency that determine good academic performance, such as cultural adjustment, motivation and the ability to cope with a variety of affective and personal factors. For Davies (2007), language proficiency has some influence on academic outcomes but personality, initiative, subject knowledge, application and motivation must also come into play and this was also seen in my cases.

The cases' personal characteristics included traits such their initiative and confidence in using English (identified in all cases). Case D showed confidence arising out of her undergraduate experiences, whilst Case C had the personality and motivation to succeed on his course despite the difficulties encountered on his written assignments, as did case B. In particular, Cases C and D both had strategies for dealing with their shortcomings in grammar or vocabulary and these "compensating strategies" are cited in other studies (Cheng et al, 2004, p.65, Andrade, 2006, p.149). Cases C and D both displayed initiative, not only in taking the lead in group presentations on account of their self-perceived better English skills, but also in their frankness in admitting to deficiencies in their use of English and their strategies in

dealing with these deficiencies without recourse to remedial English classes. For example, Case C read and wrote in his spare time, debated with others on non-academic topics and joined a scrabble club whilst D made use of English grammar and vocabulary books with which to improve her language skills. Case D was also a practising businesswoman (which related immediately to her MSc) while Case C had had some experience of his industry in an industrial safety course and had confidence in his subject area.

Bayliss and Ingram mentioned affective characteristics of individuals that could be variable in studies, such as: personality, motivation and confidence (2007, p.59). In particular, the cases in my study had the communicative confidence and ability to cope on their courses, something also alluded to in Zheng & Wei's 2014 study where Indian students, on account of their familiarity with English, had more linguistic confidence in spoken English whilst the authors identified a classroom anxiety and speaking anxiety among the Chinese students. This was also revealed in Case C and D's description of their expanding circle classmates in group work. For my outer circle West African cases, confidence in speaking was detectable in the themes from the tutorials and any anxiety was down to affective factors such as shyness (Case A) or from being away from education for a time (Case D) and not from a lack of linguistic proficiency in English.

As mentioned in the literature (Norton, 1997, Higgins, 2003, Gupta, 2006, Canagarajah, 2006, Rubdy et al, 2008), outer circle students are likely to have a language background in common with or equal to the inner circle in terms of being "legitimate speakers" (Bourdieu, 1977) of English which can aid in the adjustment to an inner circle academic environment. For example, my cases did not express any problems in integrating into the social life of the UK campus, perhaps due to their shared ownership of English with students from the inner

circle and Commonwealth "It would appear that historical, cultural and linguistic proximity play a role in helping certain nationality groups to make friends with UK students (Australia, New Zealand, USA and Canada; South Africa and Zimbabwe)" (Spencer-Oatey, Dauber, & Williams, 2014, pp. 25-26) and this aspect is reflected in other reports on international student experience in the UK "North America and Sub-Saharan Africa were more likely to have UK friends" (UKCOSA, 2004, p.12). The conclusions made in these reports reflect a kind of shared social "cultural capital" (Bourdieu, 1986) between English-speaking countries.

The cases did not mention a lack of social integration in their campuses but they did encounter and have to adjust to other varieties of English, both local and international – something anticipated in the Pearson *Relevant* factsheet (Pearson, 2012e). Variables such as encountering other Englishes, dissatisfaction with the institution, and adjusting to the new academic culture of the UK were experienced by the case study participants in their first year of their courses. This adjustment reflected findings in other studies and included becoming more organised and dealing with workload (Hill et al, 1999: Banerjee, 2003; Ferguson & White, 1994; Lowe & Cook, 2003), and generally becoming aware of what was expected of them as post-graduate students in terms of writing and working independently (Woodrow, 2006; Byrne & Flood, 2005; Hill et al, 1999; Fox, 2004). Overall, I would say that the cases were all able to perform tasks and interact in academia without linguistic proficiency being a limiting factor in those behaviours. That is not to say that they were all expert in academic skills such as applying critical thinking or organising their study time, indeed, they were sometimes lacking in some of those areas, yet linguistic proficiency was not judged to be a barrier for them in acquiring these skills.

Summary

As with Banerjee's study, there are questions raised as to the generalisability of this study's findings as I only had four cases with which to measure academic performance. However, in Banerjee's study she had categorised each participant in one of three ways: a "clear accept" "safe bet" and "risk" (2003, p.215-228, p.360) according to students' likelihood of coping with their courses, taking into account the study experiences and backgrounds of individuals and not just their test scores. All the cases in my study were examples of "clear accepts" based on:

- the inferences made about their proficiency from their PTEA score profiles
- the evidence of their proficiency in the "criterion domain" of university (expressed in their experiencing of academia)
- their ownership of English, their "legitimate speaker" status, and associated proficiency in their "native user" status (their previous academic and work experience in English)

There was a caveat that Case D's score profile did not match her apparent proficiency and confidence with English. Her confidence was high, not only from her recent graduation from her BSc and her acquisition of academic skills, but also her confidence in her subject matter which was directly related to her profession. Again, this categorisation of my cases does not mean that they did not experience performance problems. Banerjee's cases had experienced problems such as "timely effort", "adjustment to British academic conventions"; "the principle of referencing"; and "the majority of study behaviours concerned with coping with

postgraduate study" (2003, p.359-360). Similarly, all the cases in my study had to adjust to these aspects of academia.

Overall, the cases can serve as examples of test takers who scored at the requisite language level for academia (B2), coping reasonably well in their disciplines but needing adjustment to being postgraduate students. The PTEA did not signal any major deficiencies for them in relation to their language proficiency for the purpose of their academic studies but did indicate some deficiencies in C and D's profiles which, to some extent, matched their language behaviour, particularly with regard to the demands of academic writing and written discourse.

Although not generalisable to every outer circle student from Nigeria and Ghana scoring at B2 level, their experiences can serve as good illustrations of what might be happening elsewhere in other contexts. Case B from Ghana had a score profile matching the mean scores for her nationality in the PTEA score data, whilst the Nigerian cases A and C had slightly higher score profiles than the mean score for their nationality suggested by the 2014 PTEA score data. Ultimately, the cases only speak to their own unique personalities, motivations and determination. However, a look at their backgrounds and proficiency as reported by PTEA scores and a comparison to the themes in common with the other data sets in this study, suggests that they may be illustrative of the academic adjustment experiences of other outer circle African students with similar outer circle backgrounds and proficiency profiles (Gichura, 2010, Karimi, 2010, Brown 2008, Hyams-Ssekasi, 2012).

5.3 Addressing of Overall Themes

The Proficiency & Ownership of OC test takers

Nationality and L1 as examples of personal characteristics (Bachman, 2004, p.156) appear to account for variations in test performance and may be construed as "construct irrelevant variance". This is the idea that a variable or test taker characteristic is influential on the variation of scores. For example, if "native speakers" take the same proficiency test as "non-native speakers", the obvious and expected high proficiency of the former will affect the difference in test performance between these two groups and any inferences made. My findings indicated that in many cases (for example outer circle African test takers, Singaporeans, Malaysians with English L1s) outer circle test takers did perform better on the PTEA than other groups. Part of the construct of the test is that it is aimed at "non-native speakers" who are going to study in an English medium environment (appendix 5). The results of statistically significant variation in mean scores on the test according to nationality and L1 then fit with the literature on outer circle proficiency in, and ownership of English (Davies, 2013; Higgins, 2003; Kachru, 1997; Lim, 2009, 2015; Widdowson, 1994).

This confirms that many test takers in my study from the outer circle are L1 speakers of English or highly proficient second language users of English (native users) with ownership of the language. It should therefore be expected that many test takers from the outer circle will exhibit a different proficiency in English than many expanding circle test takers on account of their ESL or "native user" status (Davies, 2013) as opposed to those with an EFL orientation towards English such as the Chinese test takers in Zheng & Wei's study (2014).

The concept of ownership itself lacks overall agreement on how it manifests itself (Seilhamer, 2015) but in my study, ownership of English is demonstrated in my findings – in the proficiency of the outer circle as articulated in the PTEA mean score data, the thematic

analysis of the cases and the outer circle survey as expressed in students' educational and social backgrounds in English and their statements on their proficiency in English. In the case of the university admissions data set, I argued that scoring high on a CSELT does indicate a certain degree of ownership of English through being educated and well-qualified in English and being recognised as such by HE institutions in the UK (a recognition of students' "native user" status). This follows Davies (2011a, 2013) who argues that individuals can reach the proficiency of an "idealised educated native speaker" through education and I posit that this is true of many outer circle students such as Singaporeans, South Africans and other OC Africans, Indians, Filipinos, Malaysians and others (see tables 24, 28, 53, 54).

Ownership is also defined by the historical inheritance of English in national education systems and in the national discourse in post-colonial societies in both the inner and outer circles as well as in definitions of personal identity and perceptions of proficiency. I wouldn't have expected my case study participants to have said something along the lines of "I own English", neither would I expect this of native speakers of the inner circle, but they all laid claim to English through being educated in English; having English as part of their personal and national identity; their personal feelings of proficiency in English; and a claiming of English as a language in which they confidently operate. For many individuals in multilingual outer circle nations such as Singapore, English is increasingly a de facto L1 at home (Bolton & Ng, 2014; Lim, 2009; Lu, 2013; Tan, 2014; Wee, 2002), and this trend is apparent in other outer circle countries (D'Souza, 2001; Giri, 2015; Kurzon, 2004; Offiong & Mensah, 2012). In the case of outer circle Africa, the cases revealed that the Ghanaian and Nigerians in my study have also appropriated English (Standard and Pidgin English) despite their stated L1s on the Pearson score profile being languages other than English. Their ownership came across in the themes from discussion of their educational backgrounds and qualifications in

English, their use of English in Nigeria as a lingua franca, and their confidence in academia particularly in spoken English.

The Predictive Validity (utility of test scores) for the Cases

As discussed above, for the case studies, the PTEA score profiles can be used to infer where the individuals might have had issues in their linguistic proficiency in academia. This was apparent for cases B, C and D for whom there was sufficient data. Those scoring in the mid 50s on the GSE (B1+) or at B2 or higher (GSE 59+) did not seem to have difficulties in areas indicated by the score categories. For Cases C and D, their score profiles revealed possible weaknesses in grammar, vocabulary and written discourse which were revealed to some extent in the thematic analysis of their tutorials and analysis of their written work. For these individuals the score profiles did seem to have some predictive utility. Additionally, the cases' high proficiency in oral and aural skills was reflected in their test scores, and this translated to their confidence in communicating in English in academic situations such as presentations and group work (Cases C and D). This oral/aural confidence may have arisen from the use of spoken English in Nigerian daily life.

However, there were also aspects of the cases' academic experience that cannot be accounted for or predicted by academic English proficiency tests such as the students' use of initiative, personality, motivation and ability to get on in their subject and apply their subject knowledge and critical thinking skills (academic literacy). Other aspects of their academic performance seemed to be associated with more "universal" academic skills such as academic essay writing skills, time management, editing and referencing skills, and structuring a piece of writing to the required standards in any discipline - all issues mentioned by various studies

on academic performance (Hill et all, 1999, Brown, 2008, Cheng et al, 2004; Weir et al, 2013; Banerjee, 2003; Fox, 2004; Woodrow, 2006 and Ferguson & White, 1994) and all aspects which are not able to be predicted from general academic English language tests. Other factors such as adaptability to new learning cultures (Cases B, C and D), working with diverse groups of English speakers (Cases A, B, C and D), and coping with unsatisfactory institutional experiences (Case A) are examples of adapting to academia and academic literacies.

5.4 Implications for University and Testing Organisations

Many universities already use their own judgement as to how to evaluate a prospective student's proficiency in English. Universities do this by taking into account the previous qualifications of a student including those relating to having already been educated in the medium of English. The findings of this study lead to the recommendation that universities continue to do so because many outer circle students have legitimacy with regards having English as a first or second language and have been extensively educated in the medium of English prior to doing an IAET such as IELTS, PTE Academic or TOEFL-iBT. Moreover, the UKVI still allows this flexibility in terms of recruiting overseas students.

With regards the making of inferences from test scores, this has to be done cautiously and with the full knowledge that many other factors can influence the actual score such as the quality of preparation for the test, familiarity with the format and actual test performance on the day under the specific test conditions. Universities should also take into account other factors such as a test taker's previous education, L1, and work experience in English. A broader assessment than a reliance on inferences from test scores is needed in assessing a student's suitability for a course of study in terms of their academic literacy as well as their

level of academic English proficiency. A concern with the above would mean that the universities would have to invest a lot more time and effort into interviewing and assessing students on a case-by-case basis.

The second strand indicates that in terms of academic literacy it should be assumed that no student should automatically be thought of as academically literate just because their language proficiency may be high. Native speakers, non-native speakers, and native users can all struggle on account of not having the skills needed to critically engage with their subjects. I cannot make generalisations about the suitability of outer circle students regarding whether they should take the PTE Academic based on the experiences of only four test takers and students from West Africa. However, their experiences show that the main problem they faced was in adjusting to the academic demands of their courses, and that this was not due primarily to deficiencies in linguistic proficiency.

As regards the testing organisations behind such tests as IELTS, PTE Academic and TOEFL-iBT, there are implications for how they regard outer circle Englishes in the content of their tests, and in deciding who they subject to their tests. In terms of how they regard outer circle Englishes, it is apparent that the inner circle Englishes still hold a very important position in test construction, simply because they have been standardised and have prestige. In terms of who does their tests, testing organisations have strong commercial considerations in choosing where to market and promote their tests. For Pearson and other testing organisations, there is a need for their test in many regions around the world, as they provide an externally-validated test of academic English which they would argue is more valid, especially in terms of comprehensively testing the four skills, than many other locally-produced English language tests. Form the point of view of testing organisations, outer circle markets such as India and

Nigeria would still be legitimate places where the test can be distributed and used. However, there are other regions in the outer circle, such as South Africa and Singapore, which have many individuals who may feel affronted at having to do tests of this nature. Therefore, from a test fairness point of view, testing organisations may have to face more questions of legitimacy in certain parts of the outer circle.

With regards policies and test fairness, testing organisations will not be primarily concerned with whether a candidate is a legitimate test taker unless this policy is set by governments regarding entry to their own HE systems. A change in policy from the UKVI and universities themselves in the direction of not insisting upon tests for outer circle students would, of course, have implications for testing organisations in terms of where they then market and promote their tests. They would lose revenue from countries such as India and Nigeria if such policies were initiated. However, a change of this kind would be highly unlikely as my findings have hinted at, which indicate that there are many regions within the outer circle who do not have exceptionally high mean scores which would indicate that many individuals in those countries are not strong native users of English. As my findings show, the picture of proficiency as painted by test scores, is highly varied within the outer circle despite significant differences being found between the outer and expanding circle groups as a whole.

In an ideal world, the policies surrounding who is obliged to take English language proficiency tests should take into account the ownership of English of outer circle students but the ambiguous position of many students from the outer circle with regards proficiency in English leads to the line having to be drawn somewhere. At the moment, all outer circle countries in Africa, Asia and the Pacific are considered legitimate test takers of IAETs as

long as nationality is considered the main criterion for deciding who should take tests of this nature.

Chapter 6: Conclusion

This study has been important in investigating the proficiency in, and associated ownership of English of an under-researched student – the outer circle student. In this study I have argued that outer circle students have a very specific inheritance of English which goes hand in hand with their proficiency in the language. This then has implications for testing organisations and universities as they assess students' academic English proficiency for university, specifically a need to respect the legitimacy of the English of outer circle students. The study shows that it is important to recognise the diversity of international student experience at university (Morrison et al, 2005, p.328), and not to uncritically assign all international students into the same category in terms of both their proficiency levels and challenges that they face at university.

The claim at the beginning of the study stated that a score on an academic English proficiency test aligned with B2 would indicate sufficient proficiency in English with which to infer a student's ability to be "successful in communication in university settings" (Pearson, 2012a, p.41) and to "participate independently in higher level language interaction…participate in academic education, in both coursework and student life" (Pearson, 2012a, p.40). The case study indicated that the cases have a particular background, ownership and proficiency in English and that they largely "cope" and communicate successfully at university. In terms of the predictive validity of the PTEA for the cases in my study, there were some indications in the score profiles of areas of deficiency (particularly Cases C and D), and there was some corresponding evidence in the domain to infer that the cases were experiencing some issues in these specific linguistic areas. However, all the cases had to adapt in some way to university life, in their disciplines, and had to go through an

academic socialisation which also meant that they had to grapple with the aspects of academic literacy expected of all students. They appeared to have the threshold language proficiency with which to engage in these "higher level" and academic forms of proficiency.

For "native user" students such as the ones in the case studies, English is not a foreign language and therefore, academic support for them would differ very little in comparison to students who are from the inner circle - still meaning that they would need academic support in their discipline which would include an engagement with academic literacies. IAETs cannot provide evidence to make inferences about the ability to cope with non-linguistic aspects of academic performance. Ultimately, many of the issues encountered by the cases are part of the "adjustment" or "transition" to academia which all students have to negotiate (Lowe & Cook, 2008, Andrade, 2006, Byrne & Flood, 2005). The implication is that thinking more critically about academic literacy is a part of preparing all students for academia.

Limitations and Future Directions

In discussing the findings of this study, the limitations of the data for the study are acknowledged. In the statistical data more information on the test takers' motivations and reasons for taking the tests would be welcome as well as a caveat in terms of how they identified their L1s. In the UK university admissions survey, the views of admissions personnel regarding the quality and validity of the outer circle CSELTs was not collected and so future validity studies could look at a comparison of CSELTs and IAETs in this regard. In the case studies, there was a lack of tutor involvement – specifically a collating of their views on outer circle students' proficiency and performance in academia, and there was also

incomplete data for cases A and D in terms of tutorials and written assignments. There was also a limitation in the fact that all the cases were postgraduate students. This means that they had already been introduced to academia at undergraduate level or acquired other further qualifications to some degree before participating in this study and had therefore acquired some familiarisation of what was needed in academia. Furthermore, they had all had previous work or HE experience in English in their home country, and in the case of D, in the UK. Therefore, it would be important to investigate the experiences of recently-graduated high school students from a variety of countries studying at undergraduate level in the UK as relative "newcomers" to higher education in the inner circle, and therefore "fresher" from their domestic education systems. A further study focusing on undergraduate students would make an interesting comparison with the study described here.

Further studies focusing on either undergraduates or postgraduates could also be developed with a more varied sample in mind that would lend itself to comparisons in terms of language background, English proficiency and academic performance of cases from various nationalities within the expanding, outer *and* inner circles. This sort of comparison of experiences would be important because ultimately, academic performance and the ability to 'cope' with the demands of academia is based on many other factors beyond a sufficient proficiency in the language of instruction which can potentially impact on students from any circle of World English.

Contribution to Knowledge

I have provided empirical evidence to indicate that there is a possible proficiency difference between the outer and expanding circle in terms of their test taking populations. I have used the theory of the native user as a legitimate speaker of English to argue that asking

individuals from outer circle countries such as Nigeria, Ghana, Singapore or The Philippines, to take IAETs in addition to their domestic qualifications and experiences may be unfair and even insulting to many individuals from the outer circle. This is because many individuals from outer circle countries may be de facto native speakers of English (native users) and have already demonstrated their ability to follow educational instruction in the medium of English. Therefore, a more nuanced method of viewing and assessing English proficiency is needed when universities assess a student's English language proficiency for university. More nuanced means a continuation of using various means of assessing a potential student's English language proficiency and background and an acknowledgement that many individuals in the outer circle are legitimate speakers of English. The study contributes towards understanding and conceptualising the proficiency of outer circle students for academia in the inner circle.

The themes from the findings do suggest a legitimate objection to being asked to do IAETs among outer circle students. Why indeed should an individual have to take such a test if they have already been schooled, and socialised, in the medium of English. The tests however, do make the claim that they are a rigorous and valid test of the proficiency that a student will need at the level of HE in a country such as the UK. Overall, the study does indicate that assessment of the English language proficiency of overseas students does need to take into account issues around the legitimacy of outer circle Englishes, test fairness, and that it is often problematic to make inferences about the academic performance of test takers based solely on interpretation of performance on academic English tests.

Bibliography

Ackermann, K., De Jong, J. H. A. L., Kilgarriff, A., & Tugwell, D. (2010). Research Summary: The Pearson International Corpus of Academic English. In: Pearson Education Ltd.

Alderson, J. C., Clapham, C., & Wall, D. (1995). *Language Test Construction and Evaluation*. Cambridge: Cambridge University Press.

Alimi, M. (2011). Botswana English: implications for English teaching and assessment. *Journal of Multilingual and Multicultural Development, 32*(4), 309-324.

Allwright, J., & Banerjee, J. (1997). *Investigating the Accuracy of Admissions Criteria: a case study in a British university CRILE Occasional Report 7*. Retrieved from Lancaster:

American Psychological Association. (1954). Technical recommendations for psychological tests and diagnostic techniques. *Psychological Bulletin, 51*(2. Pt. 2.), 199-238.

Anderson, J. A. (2009). Codifying Ghanaian English: Problems and Prospects. In T. Hoffman & L. Siebers (Eds.), *World Englishes: Problems, Properties and Prospects, Selected Papers from the 13th IAWE Conference* (pp. 19-38). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Andrade, M. S. (2006). International students in English-speaking universities: Adjustment factors. *Journal of Research in International Education, 5*(2), 131-154.

Bachman, L. F. (1990). *Fundamental Considerations in Language Testing*: OUP.

Bachman, L. F. (2000). What, if any, are the limits of our responsibility for fairness in language testing? In A. J. Kunnan (Ed.), *Fairness and Validation in Language Assessment: Selected Papers from the 19th Language Testing Research Colloquium, Orlando Florida (2000)* (pp. 39-41). Cambridge, UK: UCLES.

Bachman, L. F. (2004). *Statistical Analyses for Language Assessment*. Cambridge: CUP.

Bachman, L. F. (2005). Building and Supporting a Case for Test Use. *Language Assessment Quarterly, 2*(1), 1-34.

Bachman, L. F., & Palmer, A. S. (1996). *Language Testing in Practice: Designing and Developing Useful Language Tests*. Oxford: Oxford University Press.

Bamgbose, A. (1998). Torn Between the Norms: innovations in world Englishes. *World Englishes, 17*(1), 1-14.

Banerjee, J. (2003). *Interpreting and Using Proficiency Test Scores.* (PhD), Lancaster University, Lancaster.

Barton, D. (1994). *Literacy: an introduction to the ecology of written language*. Oxford: Blackwell.

Baxter, P., & Jack, S. (2008). Qualitative Case Study Methodology: Study Design and Implementation for Novice Researchers. *The Qualitative Report, 13*(4), 544-559.

Bayliss, A., & Ingram, D. E. (2006). *IELTS as a Predictor of Academic Language Performance*. Paper presented at the Australian International Education Conference, Australian International Education Conference, Perth, Western Australia.

Baynham, M. (1995). *Literacy Practices*. London: Longman.

Bazeley, P., & Jackson, K. (2013). *Qualitative data analysis with NVivo* (2nd ed.). Los Angeles ; London: SAGE.

Becker, L. A. (2016). Effect Size Calculators. Retrieved from http://www.uccs.edu/~lbecker/ [accessed 07 June 2016]

Bolton, K., & Ng, B. C. (2014). The Dynamics of Multilingualism in Contemporary Singapore. *World Englishes, 33*(3), 307-318.

Bourdieu, P. (1973). Cultural Reproduction and Social Reproduction. In P. Brown (Ed.), *Knowledge, Education and Social Change: Papers in the Sociology of Education* (pp. 71-112). UK: Tavistock Publications.
Bourdieu, P. (1977). The Economics of Linguistic Exchanges. *Social Science Information, 16*(6), 645-668.
Bourdieu, P. (1986). The Forms of Capital. In J. Richardson (Ed.), *Handbook of Theory and Research for the Sociology of Education* (pp. 241-258). New York: Greenwood.
Braine, G. (2002). Academic literacy and the non-native speaker graduate student. *Journal of English for Academic Purposes, 1*(1), 59-68.
Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101.
Bridgeman, B., Cho, Y., & DiPietro, S. (2015). Predicting grades from an English language assessment: The importance of peeling the onion. *Language Testing*. doi:10.1177/0265532215583066
British Council. (2013). *Multilingual Education in Africa: Lessons from the Juba Language-in-Education Conference* (H. McIllwraith Ed.). London, UK.
Brown, L. (2008). The Incidence of Study-Related Stress in International Students in the Initial Stage of the International Sojourn. *Journal of Studies in International Education, 12*(1), 5-28.
Brown, L., & Holloway, I. (2008). The adjustment journey of international students at an English university. An ethnographic study. *Journal of Research in International Education, 7*(2), 232-249.
Bruthiaux, P. (2003). Squaring the circles: Issues in modelling English worldwide. *International Journal of Applied Linguistics, 13*(2).
Bryman, A. (2006). Integrating Quantitative and Qualitative Research: how is it done? *Qualitative Research, 6*(1), 97-113.
Bwenge, C. (2012). English in Tanzania: a linguistic cultural perspective. *IJLTIC, 1*(1), 167-182.
Byrne, M., & Flood, B. (2005). A study of accounting students' motives, expectations and preparedness for higher education. *Journal of Further and Higher Education, 29*(2), 111-124.
Canagarajah, S. (1999). *Resisting Linguistic Imperialism in English Teaching*. Oxford: OUP.
Canagarajah, S. (2006). Changing Communicative Needs, Revised Assessment Objectives: Testing English as an International Language. *Language Assessment Quarterly, 3*(3), 229-242.
Canale, M., & Swain, M. (1980). Theoretical Bases of Communicative Approaches to Second Language Teaching and Testing. *Applied Linguistics, 1*(1), 1-47.
Cazden, C., Cope, D., Fairclough, N., & Gee, J. (1996). A Pedagogy of Multiliteracies: designing social futures. *Havard Educational Review, 66*(1), 60-92.
Chapelle, C. A. (2012). Conceptions of Validity. In G. Fulcher & F. Davidson (Eds.), *The Routledge Handbook of Language Testing* (pp. 21-33). Abingdon, Oxon & New York: Routledge.
Chee, S. P. (2009). *Beyond the Three Circles: A New Model for World Englishes.* (Master of Arts), National University of Singapore, http://scholarbank.nus.edu.sg/bitstream/handle/10635/16502/CheeSP.PDF?sequence=1 [accessed 23 August 2016].
Cheng, L., Myles, J., & Curtis, A. (2004). Targeting Language Support for Non-Native English-Speaking Graduate Students at a Canadian University. *TESL Canada Journal, 21*(2), 50-71.

Cho, Y., & Bridgeman, B. (2012). Relationship of TOEFL iBT Scores to Academic Performance: Some evidence from American universities. *Language Testing, 29*(3), 421-442.
Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, Massachusetts: The M.I.T. Press.
Cohen, L., Manion, L., & Morrison, K. (2011). *Research Methods in Education* (7th ed.): Routledge.
Coleman, D., Starfield, S., & Hagan, A. (2003). The attitudes of IELTS stakeholders: student and staff perceptions of IELTS in Australian, UK and Chinese tertiary institutions. *IELTS Research Reports, 5*.
Cotton, F., & Conrow, F. (1998). An investigation into the predictive validity of IELTS amongst a group of international students studying at The University of Tasmania. In (Vol. 1, pp. 72-115). Canberra: IELTS Research Reports.
Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment - Structured overview of all CEFR scales*. https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=090000168045b15e.[accessed 20 August 2016]
Cramer, D., & Howitt, D. (2004). *The Sage Dictionary of Statistics*. London, Thousand Oaks California, New Delhi: Sage.
Creswell, J. W. (2009). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (3rd ed.). Thousand Oaks, CA: Sage.
Creswell, J. W. (2014). *Research Design: Qualitative, Quantitative, Mixed Methods Approaches* (4th, International Student Edition ed.). Los Angeles, Thousand Oaks, California: Sage.
Creswell, J. W., & Plano-Clark, V. L. (2011). *Designing and Conducting Mixed Methods Research* (2nd ed.). Los Angeles, London: SAGE Publications.
Cronbach, L. J. (1990). *Essentials of Psychological Testing* (5th ed.). New York: Harper & Row.
Cronbach, L. J., & Meehl, P. E. (1955). Construct Validity in Psychological Tests. *Psycholgical Bulletin, 52*(4), 281-302.
Crystal, D. (2003). *English as a Global Language* (Second ed.). Cambridge: Cambridge University Press.
Crystal, D. (2012). A Global Language. In P. Seargeant & J. Swann (Eds.), *English in the world: history, diversity, change* (pp. 152-177): Routledge.
Cummins, J. (1979). Cognitive/academic language proficiency, linguistic interdependence, the optimum age question and some other matters. *Working Papers on Bilingualism, 19*, 121-129.
Cummins, J. (2008). BICS and CALP: Empirical and Theoretical Status of the Distinction. In B. Street & N. H. Hornberger (Eds.), *Encyclopedia of Language and Education* (2nd ed., Vol. 2, pp. 71-83). New York: Springer Science + Business Media LLC.
D'Souza, J. (2001). Contextualizing Range and Depth in Indian English. *World Englishes, 20*(2), 145-159.
Davies, A. (2007). Assessing Academic English Language Proficiency: 40+ years of UK language tests. In J. Fox, M. Wesche, D. Bayliss, L. Cheng, C. Turner, & C. Doe (Eds.), *Language Testing Reconsidered* (pp. 73-86). Ottowa, Canada: University of Ottawa Press.
Davies, A. (2011a). Does Language Testing Need the Native Speaker? *Language Assessment Quarterly, 8*(3), 291-308.
Davies, A. (2011b). Kane, validity and soundness. *Language Testing, 29*(1), 37-42.

Davies, A. (2013). *Native Speakers and Native Users: loss and gain*. Cambridge: Cambridge University Press.
Davies, A., Brown, A., Elder, C., Hill, K., Lumley, T., & McNamara, T. (1999). *Dictionary of Language Testing*. Cambridge: UCLES/CUP.
Davies, A., Hamps-Lyons, L., & Kemp, C. (2003). Forum - Whose norms? International proficiency tests in English. *World Englishes, 22*(4), 571-584.
De Jong, J. H. A. L., & Zheng, Y. (2011). Research Note: Applying EALTA Guidelines - A practical case study on Pearson Test of English Academic. In (pp. 15): Pearson Education Ltd.
De Klerk, V., & Bosch, B. (1998). Language Shift in Post-apartheid South Africa: the appeal of English. *LACUS Forum, 24*, 353-363.
Dooey, P., & Oliver, R. (2002). An investigation into the predictive validity of the IELTS test. *Macquarie University Prospect Journal, 17*(1), 36-54.
Duff, A., Boyle, E., Dunleavy, K., & Ferguson, J. (2004). The relationship between personality, approach to learning and academic performance. *Personality and Individual Differences, 36*(8), 1907-1920.
Duff, P. A. (2008). *Case Study Research in Applied Linguistics*. New York: Lawrence Erlbaum.
Educational Testing Service. (2010). *Linking TOEFL-iBT Scores to IELTS Scores - A Research Report*. Retrieved from https://www.ets.org/s/toefl/pdf/linking_toefl_ibt_scores_to_ielts_scores.pdf: [accessed 07 June 2016]
Educational Testing Service. (2014). Test and Score Data Summary for TOEFL-iBT Tests: January 2013 - December 2013 Test Data. Retrieved from https://www.ets.org/s/toefl/pdf/94227_unlweb.pdf [accessed 14 May 2015]
Educational Testing Service. (2015). Compare TOEFL Scores. Retrieved from https://www.ets.org/toefl/institutions/scores/compare/ [accessed 14 May 2015]
Edwards, V., Ran, A., & Li, D. (2007). Uneven playing field or falling standards?: Chinese students' competence in English. *Race, Ethnicity and Education, 10*(4), 387-400.
Elder, C. (1993). Language proficiency as a predictor of performance in teacher education. *Melbourne Papers in Language Testing, 2*(1), 72-95.
English, F. (1999). What Do Students Really Say In Their Essays? In C. Jones, J. Turner, & B. Street (Eds.), *Students Writing in the University: cultural and epistemological isssues*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
Feast, V. (2002). The Impact of IELTS Scores on Performance at University. *International Education Journal, 3*(4), 70-85.
Ferguson, G., & White, E. (1992). A small-scale study of predictive validity. *Melbourne Papers in Language Testing*, 15-63.
Ferguson, G., & White, E. (1994). A Predictive Validity Study of IELTS. In (pp. 15-63). Institute of Applied Linguistics, University of Edinburgh: *Melbourne Papers in Language Testing* IELTS Research Reports.
Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics* (4th ed.). London, Thousand Oaks: Sage.
Fox, J. (2004). Test decisions over time: Tracking validity. *Language Testing, 21*(4), 437-465.
Freeman, D. (2015). *Frozen in Thought? How we think and what we do in ELT*. Paper presented at the IATEFL 2015, Manchester, UK.
Fulcher, G., & Davidson, F. (2007). *Language Testing and Assessment*. London & New York: Routledge.

Galloway, N., & Rose, H. (2015). *Introducing Global Englishes* (1st ed.). London, New York: Routledge.
Gibbons, P. (2009). *English Learners, Academic Literacy, and Thinking: Learning in the Challenge Zone*. Portsmouth, NH, USA: Heinemann Educational Books.
Gichura, S. W. (2010). *Experience and Study Outcomes of Kenyan Students Pursuing Higher Education in the UK.* (PhD Thesis), Coventry University.
Gill, S. K. (2005). Language Policy in Malaysia: Reversing Direction. *Language Policy, 4,* 241-260
Gilliver-Brown, K., & Johnson, E. M. (2009). *Academic Literacy Development: A Multiple Perspectives Approach to Blended Learning.* Paper presented at the ASCILITE, Auckland, New Zealand.
Giri, R. A. (2015). The many faces of English in Nepal. *Asian Englishes, 17*(2), 94-115. doi:http://dx.doi.org/10.1080/13488678.2015.1003452
Goertz, G. (2006). *Social Science Concepts: A user's guide*. Princeton, N.J.: Princeton University Press.
Gorlach, M. (1988). The Development of Standard English. In M. Gorlach (Ed.), *Studies in the History of the English Language* (pp. 9-64). Heidlberg: Carl Winter.
Graddol, D. (2006). *English Next*. UK: British Council.
Graham, J. G. (1987). English Language Proficiency and the Prediction of Academic Success. *TESOL Quarterly, 21*(3), 505-521.
Greene, J. C. (2007). Mixed Methods in Social Inquiry.
Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a Conceptual Framework for Mixed-Method Evaluation Designs. *Educational Evaluation and Policy Analysis, 11*(3), 255-274.
Guardian News and Media Limited. (2015). University league table 2015 - the complete list. Retrieved from http://www.theguardian.com/education/ng-interactive/2014/jun/02/university-league-tables-2015-the-complete-list [accessed 13 October 2015]
Gupta, A. F. (2006). Standard English in the World. In R. Rubdy & M. Saraceni (Eds.), *English in the World: Global Rules, Global Roles* (pp. 95-109). London & New York: Continuum.
Hamps-Lyons, L., & Davies, A. (2008). The Englishes of English tests: bias revisited. *World Englishes, 27*(1), 26-39.
Hashim, A. (2014). English and the Linguistic Ecology of Malaysia. *World Englishes, 33*(4), 458-471.
Havergal, C. (2015). Glyndwr faces continued student visa restrictions after review. *Times Higher Education*. Retrieved from https://www.timeshighereducation.com/news/glyndwr-faces-continued-student-visa-restrictions-after-review/2018881.article [accessed 20 February 2016]
Higgins, C. (2003). "Ownership" of English in the Outer Circle: An Alternative to the NS-NNS Dichotomy. *TESOL QUARTERLY, 37*(4, Winter), 615-644.
Higgins, C. (2009). *English as a Local Language: Post-colonial identities and multilingual practices*. Bristol, Buffalo, Toronto: Multilingual Matters.
Hill, K., Storch, N., & Lynch, B. (1999). A comparison of IELTS and TOEFL as predictors of academic success. *IELTS Research Reports, 2*, 62-73.
Holm, J. (1983). *Central American English* (J. Holm Ed.). Heidelberg: Julius Groos Verlag Heidelberg.
Howell, B., Nathan, P., Schmitt, D., Sinclair, C., Spencer, J., & Wrigglesworth, J. (2012). BALEAP Guidelines on English Language Tests for University Entry. In: BALEAP.

Hulstijn, J. H. (2011). Language Proficiency in Native and Nonnative speakers: An agenda for research and suggestions for second-language assessment. *Language Assessment Quarterly, 8*(3), 229-249.
Hyams-Ssekasi, D. (2012). *The Transition of International Sub-Saharan African Students into the UK University system with reference to a University in the North of England.* (PhD Thesis), University of Huddersfield, http://eprints.hud.ac.uk/16428/ [accessed 20 August 2016]
Hyatt, D., & Brooks, G. (2009). Investigating stakeholders' perceptions of IELTS as an entry requirement for higher education in the UK. *IELTS Research Reports, 10*.
Hymes, D. H. (1972). On Communicatie Competence. In J. B. Pride & J. Holmes (Eds.), *Sociolingustics* (pp. 269-293). Harmondsworth: Penguin.
IELTS. (2016). Common European Framework. Retrieved from https://www.ielts.org/ielts-for-organisations/common-european-framework [accessed 24 June 2016]
IELTS Partners. (2013). IELTS Researchers - Test Taker Performance 2013. Retrieved from http://www.ielts.org/researchers/analysis_of_test_data/test_taker_performance_2013.aspx [accessed 12 March 2015]
Igboanusi, H., & Wolf, H. G. (2009). the role of ethnically mixed marriages in language shift: a case study of Nigeria's minority languages. *Sociolinguistic Studies, 3*(3). doi:10.1558/sols.v3i3.451
Ingram, D. E., & Bayliss, A. (2007). IELTS as a predictor for academic language performance Part 1. *IELTS Research Reports, 7*.
Jenkins, J. (2006). The Spread of EIL: A testing time for testers. *ELT Journal, 60*(1), 42-50.
Jenkins, J. (2009). *World Englishes: a source book for students* (2nd ed.). London and New York: Routledge.
Jenkins, J. (2014). Who Speaks English Today? In J. Jenkins (Ed.), *Global Englishes: A Resource Book for Students* (pp. 10-20). London: Routledge.
Jick, T. D. (1979). Mixing Qualitative and Quantitative Methods: Triangulation in Action. *Administrative Science Quarterly, 24*(4), 602-611.
Johnson, P. (1988). English language proficiency and academic performance of undergraduate international students. *TESOL Quarterly, 22*(1), 164-168.
Johnson, R. B., & Onuwuegbuzie, A. J. (2004). Mixed Methods Research: A Paradigm Whose Time Has Come. *Educational Researcher, 33*(7), 14-26.
Kachru, B. B. (1985). Standards, codification and sociolinguistic realism: The English language in the Outer Circle. In R. Quirk & H. G. Widdowson (Eds.), *English in the World: Teaching and learning the language and literatures* (pp. 11-30). Cambridge: Cambridge University Press.
Kachru, B. B. (1986). *The Alchemy of English: the Spread, Functions and Models of Non-native Englishes*. Oxford: Pergamon Institute of English.
Kachru, B. B. (1992). *The Other Tongue: English across Cultures* (2nd ed.). Chicago: University of Illinois Press.
Kachru, B. B. (1996). World Englishes: Agony and Ecstasy. *Journal of Aesthetic Education, 30*(2), 135-155.
Kachru, B. B. (1997). World Englishes and English-Using Communities. *Annual Review of Applied Linguistics, 17*, 66-87.
Kachru, B. B. (1998). English as an Asian Language. *Links and Letters, 5*, 89-108.
Kachru, B. B., Kachru, Y., & Nelson, C. L. (2006). *The Handbook of World Englishes*. Malden, MA ; Oxford: Blackwell.
Kamwangamalu, N. (2009). South African Englishes. In B. B. Kachru, Y. Kachru, & C. Nelson (Eds.), *The Handbook of World Englishes* (pp. 158-171): Wiley-Blackwell.

Kamwangamalu, N., & Tovares, A. (2016). English in language ideologies, attitudes and educational practices in Kenya and South Africa. *World Englishes*. doi:10.1111/weng.12207
Kamwendo, G. H. (2003). Is Malawi Guilty of Spoiling the Queen's Language? *English Today 74, 19*(2), 30-33.
Kane, M. T. (1994). Validating the Performance Standards Associated with Passing Scores. *Review of Educational Research, 64*(3), 425-461.
Kane, M. T. (2011). Validating Score Interpretations and Uses: Messick Lecture, Language Testing Research Colloquium, Cambridge, April 2010. *Language Testing, 29*(1), 3-17.
Karimi, F. K. (2010). Factors Contributing to the Academic Performance of Students in a Private University in Kenya: A Case of Daystar University. *International Journal of Learning, 17*(8), 63-78.
Kerstjens, M., & Nery, C. (2000). Predictive validity in the IELTS Test: a study of the relationship between IELTS scores and students' subsequent academic performance. *IELTS Research Reports, 3*.
Khan, S. Z. (2009). Imperialism of International Tests: an EIL perspective. In F. Sharifian (Ed.), *English as an International Language - Perspectives and Pedagogical Issues* (pp. 190-205). Bristol, UK: Multilingual Matters.
Kish, L. (1959). Some Statistical Problems in Research Design. *American Sociological Review, 24*(3), 328-338.
Kruger, J., & Dunning, D. (2009). Unskilled and Unaware of it: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-Assessments. *Psychology, 1*, 30-46.
Kunnan, A. J. (2000). Fairness and Justice for All. In A. J. Kunnan (Ed.), *Fairness and Validation in Language Assessment: Selected Papers from the 19th Language Testing Research Colloquium, Orlando Florida (2000)* (pp. 1-14). Cambridge: UCLES/CUP.
Kunnan, A. J. (2010). Test fairness and Toulmin's argument structure. *Language Testing, 27(2)*, 183-189.
Kurzon, D. (2004). *Where East Looks West - Success in English in Goa and on the Konkan Coast*. Clevedon, Buffalo, Toronto, Sydney: Multilingual Matters.
Landsberger, H. A. (1958). *Hawthorne Revisited, .* New York: Ithaca.
Lea, M. R., & Street, B. R. (1998). Student Writing in Higher Education: An academic literacies approach. *Studies in Higher Education, 23*(2), 157-172.
Leung, C., Harris, R., & Rampton, B. (1997). The Idealised Native Speaker, Reified Ethnicities and Classroom Realities. *TESOL Quarterly, 31*(3), 543-560.
Li, S., Marquart, J. M., & Zercher, C. (2000). Conceptual Issues and Analytic Strategies in Mixed-Methods Studies of Pre-school Inclusion. *Journal of Early Intervention, 23*(2), 116-132.
Light, R. L., Xu, M., & Mossop, J. (1987). English proficiency and academic performance of international students. *TESOL Quarterly, 21*, 251-261.
Lim, L. (2009). Beyond Fear and Loathing in SG: The real mother tongues and language policies in multilingual Singapore. *AILA Review, 22*, 52-71.
Lim, L. (2015). Coming of age, coming full circle: the (re)positioning of (Singapore) English and multilingualism in Singapore at 50. *Asian Englishes, 17*(3), 261-270.
Linn, R. L. (1990). Admissions Testing: Recommended Uses, Validity, Differential Prediction and Coaching. *Applied Measurement in Education, 3*(4), 297-318.
Lloyd-Jones, G., Neame, C., & Medaney, S. (2012). A multiple case study of the relationship between the indicators of students' English language competence on entry and

students' academic progress at an international postgraduate university. *IELTS Research Reports, 11*, 1-54.
Lowe, H., & Cook, A. (2003). Mind the Gap: Are students prepared for higher education? *Journal of Further and Higher Education, 27*(1), 53-76.
Lowenberg, P. H. (1986). Non-native Varieties of English: Nativization, Norms and Implications. *SSLA, 8*, 1-18.
Lowenberg, P. H. (1993). Issues of Validity in Tests of English as a World Language: Whose norms? *World Englishes, 12*(1), 95-106.
Lowenberg, P. H. (2000). Non-native varieties and issues of fairness in testing English as a world language. In A. J. Kunnan (Ed.), *Studies in Language Testing 9: Fairness and Validation in Language Assessment* (pp. 43-59): CUP.
Lu, L. (2013). Can English be a Singaporean Mother Tongue? *Today*. Retrieved from www.todayonline.com/commentary/can-english-be-singaporean-mother-tongue [accessed 12 December 2015]
Macmillan. (2007). *Macmillan English Dictionary for Advanced Learners* (2nd ed.). UK: Macmillan.
Matiki, A. J. (2001). The Social Significance of English in Malawi. *World Englishes, 20*(2), 201-218.
Mazrui, A. M. (2004). *English in Africa After the Cold War*. Clevedon, England: Multilingual Matters Ltd.
McArthur, T. (1987). The English Languages? *English Today, 3*(3), 9-13.
McGrath, R. E., & Meyer, G. J. (2006). When Effect Sizes Disagree: the Case of *r* and *d*. *Psychological Methods, 11*(4), 386-401.
McNamara, T. F. (2003). Looking back, looking forward: rethinking Bachman. *Language Testing, 20*(4), 466-473.
McNamara, T. F. (2006). Validity in Language Testing: The Challenge of Sam Messick's Legacy. *Language Assessment Quarterly, 3*(1), 31-51.
Messick, S. (1980). Test validity and the ethics of assessment. *American psychologist, 35*(11), 1012.
Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational Measurement* (3rd ed., pp. 13-104). New York: American Council on Education and Macmillan.
Mesthrie, R. (1992). *English in Language Shift: The history, structure and sociolinguistics of South African Indian English*. Cambridge, New York: Cambridge University Press.
Michieka, M. M. (2005). English in Kenya: a sociolinguistic profile. *World Englishes, 24*(2), 173-186.
Michieka, M. M. (2009). Expanding Circles within the Outer Circle: The rural Kisii in Kenya. *World Englishes, 28*(3), 352-364.
Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis : an expanded sourcebook* (2nd ed.). Thousand Oaks, Calif. ; London: Sage.
Miles, M. B., Huberman, A. M., & Saldana, J. (2014). *Qualitative Data Analysis: A Methods Sourcebook* (3rd ed.): Sage.
Modiano, M. (1999). International English in the Global Village. *English Today, 58*(2), 22-28.
Monash University. (2014). Grade Point Average (GPA). Retrieved from http://www.monash.edu.au/exams/gpa.html [accessed 03 May 2014]
Morrison, J., Merrick, B., Higgs, S., & Le Metais, J. (2005). Researching the performance of international students in the UK. *Studies in Higher Education, 30*(3), 327-337.
Morse, J. M. (1991). Approaches to Quanlitative-Quantitative Methodological Triangulation. *Nursing Research, 40*, 120-123.
Munby, J. (1978). *Communicative Syllabus Design*. Cambridge: Cambridge University Press.

Murray, N. (2010). Conceptualising the English language needs of first year university students. *The International Journal of the First Year in Higher Education, 1*(1), 55-64.
Nesi, H. (2012). Laughter in University Lectures. *Journal of English for Academic Purposes, 11*, 79-89.
Norton, B. (1997). Language, Identity and The Ownership of English. *TESOL Quarterly, 31*(3), 409-429.
Nunan, D. (2009). *Second Language Teaching and Learning*. Boston: Heinle & Heinle.
O'Hara-Davies, B. (2010). The Paradox of English. *Journal of Multilingual and Multicultural Development, 31*(2), 107-118.
O'Loughlin, K. (2008). The Use of IELTS for University Selection in Australia: A case study. *IELTS Research Reports, 2008*.
O'Sullivan, B. (2000). *Towards a Model of Performance in Oral Language Testing.* (PhD Thesis), University of Reading.
Obaidul Hamid, M. (2014). World Englishes in International Proficiency Tests. *World Englishes, 33*(2), 263–277.
Ochieng, D. (2015). The Revivial of the Status of English in Tanzania. *English Today, 31*(2), 25-31.
Offiong, O. A., & Mensah, E. O. (2012). Language Choice and Family Language Policy in Inter-Ethnic Marriages in South-Eastern Nigeria. *Studies in Literature and Language, 4*(2).
OUP. (2002). *Concise Oxford English Dictionary* (J. Pearsall Ed. 10th ed.). New York: OUP.
Pakir, A. (1997). Standards and Codification for World Englishes. In L. E. Smith & M. L. Forman (Eds.), *Literary Studies East and West, World Englishes 2000, Selected Essays* (Vol. 14). University of Hawai'i: University of Hawai'i.
Pallant, J. (2010). *SPSS Survival Manual - A step by step guide to data analysis using SPSS* (4th ed.). Maidenhead, New York: McGraw Hill.
Patton, M. Q. (1990). *Qualitative evaluation and research methods* (2nd ed.). Newbury Park, Calif.: Sage Publications.
Pearson (2010a). Aligning PTE Academic Test Scores to the Common European Framework of Reference for Languages. Pearson Education Ltd.
Pearson (2010b). *The Official Guide to PTE: Pearson Test of English Academic*. Hong Kong: Pearson Education Asia Ltd.
Pearson (2012a). PTE Academic Score Guide. Pearson Education Ltd.
Pearson (2012b). Interpreting the PTE Academic Score Report. Pearson Education Ltd.
Pearson (2012c). Validity and Reliability in PTE Academic. Pearson Education Ltd.
Pearson (2012d). PTE Academic Test Taker Handbook. Pearson Education Ltd.
Pearson (2012e). Relevant Factsheet. Retrieved from http://pearsonpte.com/wp-content/uploads/2014/07/RelevantFactsheet.pdf [accessed 09 September 2014]
Pearson (2015). PTE Academic, UK Higher Education Institutions Information Pack, prepared June 2015. Harlow, Essex UK: Pearson.
Pierson, H. D. (1998). Societal Accommodation to English and Putonghua in Cantonese-Speaking Hong Kong. In M. C. Pennington (Ed.), *Language in Hong Kong at Century's End* (pp. 91-112). Hong Kong: Hong Kong University Press.
Read, J. (2015). *Assessing English Proficiency for University Study*. London & New York: Palgrave Macmillan.
Riazi, M. (2013). Concurrent and Predictive Validity of Pearson Test of English Academic (PTE Academic). *Papers in Language Testing and Assessment, 2*(2), 1-27.
Riessman, C. K. (1993). *Narrative Analysis*. Newbury Park, London, New Delhi: Sage.

Rowntree, D. (1991). *Statistics Without Tears: an introduction for non-mathematicians*: Penguin.
Rubdy, R., McKay, S., Alsagoff, L., & Bokhorst-Heng, W. D. (2008). Enacting English language ownership in the Outer Circle: a study of Singaporean Indians' orientations towards English norms. *World Englishes, 27*(1), 40-67.
Salvadori, C. (1983). *Through Open Doors: A View of Asian Cultures in Kenya* (A. Fedders Ed.). Nairobi, Kenya: Kenway Publications Ltd.
Schmied, J. (1991). *English in Africa: An introduction*. London and New York: Longman.
Schmitz, J. R. (2012). "To ELF or not to ELF?" (English as a Lingua Franca): That's the question for Applied Linguistics in a globalised world. *Rev. Bras. linguist. apl., 12*(2), 249-284.
Schneider, E. W. (2007). *Postcolonial English: Varieties around the world*. UK: Cambridge University Press.
Seidlhofer, B. (2005). Key Concepts in ELT: English as a Lingua Franca. *ELT Journal, 59*(4), 339-341.
Seilhamer, M. F. (2015). The Ownership of English in Taiwan. *World Englishes, 34*(3), 370-388.
Shaw, S. D., & Weir, C. J. (2007). *Examining Writing: Research and Practice in Assessing Second Language Writing*. Cambridge: Cambridge University Press.
Shepherd, J. (2013). Ban on London Met University enrolling foreign students lifted. *The Guardian*. Retrieved from http://www.theguardian.com/education/2013/apr/09/ban-university-foreign-students-lifted [accessed 24 August 2016]
Spack, R. (1988). Initiating ESL Students into the Academic Discourse Community: How Far Should We Go? *TESOL Quarterly, 22*(1), 29-51.
Spack, R. (1997). The Acquisition of Academic Literacy in a Second Language: A longitudinal case study. *Written Communication, 14*(1), 3-62.
Spencer, J. (1966). The Anglo-Indians and Their Speech: A Sociolinguistic Essay. *Lingua, 16*, 57-70.
Spencer-Oatey, H., Dauber, D., & Williams, S. (2014). *Promoting Integration on Campus: Principles, Practice and Issues for Further Exploration*. Retrieved from http://institutions.ukcisa.org.uk/Info-for-universities-colleges--schools/Publications--research/resources: http://institutions.ukcisa.org.uk//Info-for-universities-colleges--schools/Publications--research/resources/70/Promoting-Integration-on-Campus-Principles-Practice-and-Issues-for-Further-Exploration [accessed 14 July 2016]
Stake, R. E. (1995). *The art of case study research*. Thousand Oaks, Calif. ; London: Sage.
Stephen, J. (2013). English in Malaysia: a case of the past that never really went away? *English Today, 29(2)*, 3-8.
Stoffelsma, L., & De Jong, J. H. A. L. (2015). The English Reading Proficiency of Future Teachers in Ghana. *Nordic Journal of African Studies, 24*(2), 94-117.
Street, B. (1984). *Literacy in Theory and Practice*. Cambridge: Cambridge University Press.
Strevens, P. (1992). English as an International Language: Directions in the 1990s. In B. B. Kachru (Ed.), *The Other Tongue: English across Cultures* (2nd ed., pp. 27-47). Urbana and Chicago: University of Illinois Press.
Swansea University. (2014). Taught via the Medium of English. Retrieved from http://www.swansea.ac.uk/postgraduate/apply/taughtviathemediumofenglish/ [accessed 20 July 2016]
Symonds, J. E., & Gorard, S. (2010). *The Death of Mixed Methods: Research Labels and their Casualites*. Paper presented at the The British Educational Research Association Annual Conference, Heriot Watt University, Edinburgh. http://www.leeds.ac.uk/educol/documents/174130.pdf [accessed 24 July 2015]

Tan, Y. Y. (2014). English as a 'mother tongue' in Singapore. *World Englishes, 33*(3), 319-339.
Tapson, F. (2006). *Oxford Mathematics Study Dictionary* (3rd ed.). Oxford: Oxford University Press.
Taylor, L. (2006). The changing landscape of English: implications for language assessment. *ELT Journal, 60*(1), 51-60.
Taylor, L. (2009). Setting Language Standards for Teaching and Assessment: a matter of principle, politics or prejudice? In L. Taylor & C. J. Weir (Eds.), *Language Testing Matters: Investigating the wider social and educational impact of assessment* (pp. 139-157). Cambridge: Cambridge University Press.
Teddlie, C., & Tashakkori, A. (2006). A General Typology of Research Designs Featuring Mixed Methods. *Research in the Schools, 13*(1), 12-28.
Teddlie, C., & Tashakkori, A. (2009). *Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences.* Thousand Oaks, CA: Sage.
The World Bank. (2015). Population Total. Retrieved from http://data.worldbank.org/indicator/SP.POP.TOTL [accessed 20 August 2016]
Thomas, G. (2011). *How to do your case study : a guide for students and researchers*. Los Angeles, Calif. ; London: SAGE.
Tokunaga, H. T. (2015). *Fundamental Statistics for the Social and Behavioral Sciences*. Thousand Oaks, California: SAGE Publications Inc.
UCAS. (2014). International Qualifications for Entry to University or College in 2015. In (pp. 102). UK: UCAS.
UCLES. (2015). Grade Statistics. Retrieved from http://www.cambridgeenglish.org/research-and-validation/quality-and-accountability/grade-statistics/ [accessed 15 May 2015]
UK Council for International Student Affairs. (2011). The UKCISA Tier 4 Student Survey 2011. In. UK.
UK Visas and Immigration. (2016a). Guidance: Applying for a UK visa: approved English language tests. 28 July 2016. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/541659/Approved_Secure_English_Language_Tests_and_Test_Centres.pdf [accessed 20 August 2016]
UK Visas and Immigration. (2016b). *Tier 4 of the Points Based System - Policy Guidance: Version 4/2016*. UK: Home Office Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/261401/Tier4migrantguidance.pdf. [accessed 20 August 2016]
Universities UK. (2015). Universities UK Members. Retrieved from http://www.universitiesuk.ac.uk/aboutus/members/Pages/default.aspx [accessed 14 July 2015]
University of Leicester. (2016). Countries where education is in English. Retrieved from https://le.ac.uk/student-life/international-students/english-language-requirements/approved-countries [accessed 18 July 2016]
University of Liverpool. (2014). English Language Entry Requirements. Retrieved from http://www.liv.ac.uk/study/international/countries/english-language/ [accessed 29 April 2014]
Verschuren, P. (2003). Case study as a research strategy: some opportunities and ambiguities. *International Journal of Social Research Methodology, 6*(2), 121-139.
WAEC. (2000). *The West African Examinations Council Annual Report 2000*. Retrieved from Lagos, Nigeria:

Wait, I. W., & Gressel, J. W. (2009). Relationship Between TOEFL Score and Academic Success for International Engineering Students. *Journal of Engineering Education, October*, 389-399.

Wang, H., Choi, I., Schimdgall, J., & Bachman, L. (2012). Review of Pearson Test of English Academic: Building an assessment use argument. *Language Testing, 29*(4), 603-619.

Wee, L. (2002). When English is not a Mother Tongue: Linguistic Ownership and the Eurasian Community in Singapore. *Journal of Multicultural and Multilingual Development, 23*(4), 282-295.

Weideman, A. (2012). Validation and Validity Beyond Messick. *Per Linguam, 28*(2), 1-14.

Weir, C. J. (2005). *Language testing and validation : an evidence-based approach*. Basingstoke ; New York: Palgrave Macmillan.

Weir, C. J., Chan, S. H. C., & Nakatsuhara, F. (2013). Examining the criterion-related validity of the GEPT advanced reading and writing tests: Comparing GEPT with IELTS and real-life academic performance. *LTTC-GEPT Research Reports RG-01*, 1-43.

Widdowson, H. G. (1994). The Ownership of English. *TESOL Quarterly, 28*(2), 377-389.

Widdowson, H. G. (1997). THE FORUM EIL, ESL, EFL: global issues and local interests. *World Englishes, 16*(1), 135-146.

Williams, E. (2011). Language Policy, Politics and Development in Africa. In H. Coleman (Ed.), *Dreams and Realities: Developing countries and the English language* (pp. 41-58). UK: British Council.

Wingate, U., & Tribble, C. (2012). The Best of Both Worlds?: Towards an English for Academic Purposes/Academic Literacies Writing Pedagogoy. *Studies in Higher Education, 37*(4), 481-495.

Woodrow, L. (2006). Academic success of international postgraduate education students and the role of English proficiency. *University of Sydney Papers in TESOL, 1*, 51-70.

Yin, R. K. (2006). Mixed Methods Research: Are the Methods Genuinely Integrated or Merely Parallel? *Research in the Schools, 13*(1), 41-47.

Yin, R. K. (2009). *Case Study Research: Design and Methods* (4th ed.). Thousand Oaks, CA: Sage.

Zheng, Y., & Wei, W. (2014). Knowing the Test Takers: Investigating Chinese and Indian (EFL/ESL) Students' Performance on PTE Academic. *The Asian EFL Journal Quarterly, 16*(1), 120-152.

Appendices

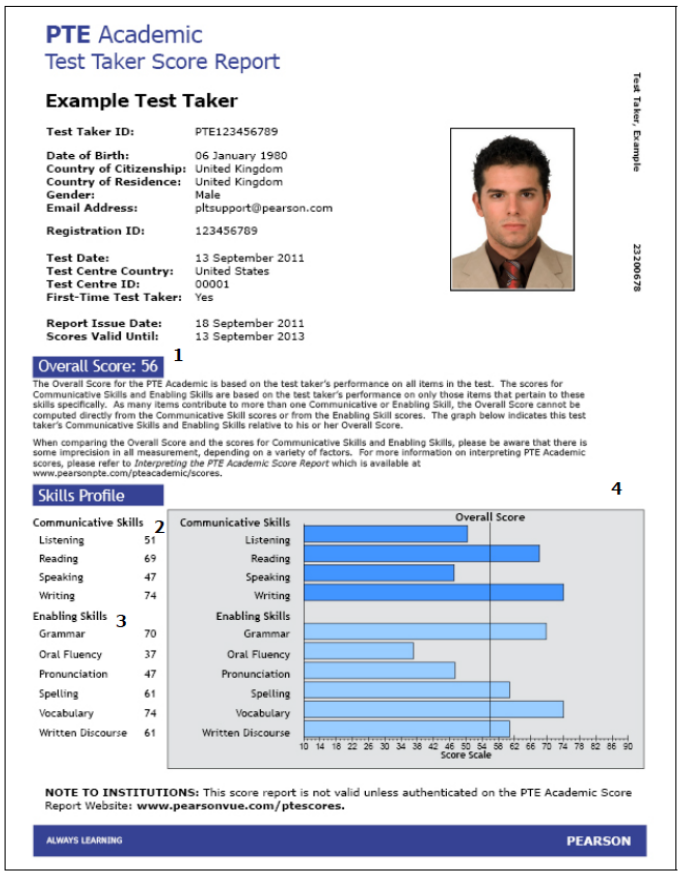
Appendix 1: The PTE Academic Score Report (Skills Profile)

Interpreting the PTE Academic Score Report

1. Get your scores

After taking the test, you will be notified by email when your PTE Academic scores are available (typically within five business days from your test date). Using the online account you set up when you registered for the test, you can view your results and then request to make them available to an unlimited number of institutions of your choice.

The PTE Academic Score Report includes your overall score, communicative skills scores and enabling skills scores.



PTE Academic
Test Taker Score Report

Example Test Taker

Test Taker, Example

Test Taker ID:	PTE123456789
Date of Birth:	06 January 1980
Country of Citizenship:	United Kingdom
Country of Residence:	United Kingdom
Gender:	Male
Email Address:	pltsupport@pearson.com
Registration ID:	123456789
Test Date:	13 September 2011
Test Centre Country:	United States
Test Centre ID:	00001
First-Time Test Taker:	Yes
Report Issue Date:	18 September 2011
Scores Valid Until:	13 September 2013

23200678

1

Overall Score: 56

The Overall Score for the PTE Academic is based on the test taker's performance on all items in the test. The scores for Communicative Skills and Enabling Skills are based on the test taker's performance on only those items that pertain to these skills specifically. As many items contribute to more than one Communicative or Enabling Skill, the Overall Score cannot be computed directly from the Communicative Skill scores or from the Enabling Skill scores. The graph below indicates this test taker's Communicative Skills and Enabling Skills relative to his or her Overall Score.

When comparing the Overall Score and the scores for Communicative Skills and Enabling Skills, please be aware that there is some imprecision in all measurement, depending on a variety of factors. For more information on interpreting PTE Academic scores, please refer to *Interpreting the PTE Academic Score Report* which is available at www.pearsonpte.com/pteacademic/scores.

4

Skills Profile

Communicative Skills	2
Listening	51
Reading	69
Speaking	47
Writing	74
Enabling Skills	3
Grammar	70
Oral Fluency	37
Pronunciation	47
Spelling	61
Vocabulary	74
Written Discourse	61

NOTE TO INSTITUTIONS: This score report is not valid unless authenticated on the PTE Academic Score Report Website: **www.pearsonvue.com/ptescores.**

ALWAYS LEARNING PEARSON

Example Test Taker Score Report

1. The **overall score** reflects your overall English language ability. The score is based on performance on all items in the test.

   The range for the overall score is 10-90 points.

2. **Scores for communicative skills** (listening, reading, speaking and writing) are based on all test items that assess these skills, either as a single skill or together with other skills.

   The range for each communicative skill score is 10-90 points.

3. **Scores for enabling skills** (grammar, oral fluency, pronunciation, spelling, vocabulary and written discourse) are based on all test items assessing one or more of these skills.

   The range for each enabling skill score is 10-90 points.

4. The display of the scores in a graph allows you to quickly see your strengths and weaknesses, and how each skill relates to your overall performance.

Source: Pearson (2012b), Interpreting the Academic Score Report, p.1
Available from the PTE Academic website as a PDF file:
http://www.pearsonlongman.com/examsplace/PTE_Academic/index.html [Accessed 10/06/2016]

Appendix 2: Majority English Speaking Countries (MESCs)

Exceptions to the English language requirement

127. You do not have to demonstrate English language competence if you:

d) have previously completed an academic qualification equivalent to a UK degree which was taught in what we call a 'majority English speaking' country; (see the table below):

A – H	I – P	Q – Z
Antigua and Barbuda	Ireland	St Kitts and Nevis
Australia	Jamaica	St Lucia
The Bahamas	New Zealand	St Vincent and the Grenadines
Barbados		Trinidad and Tobago
Belize		United States of America
Dominica		
Grenada		
Guyana		

e) are a national from what we class as 'majority English- speaking' countries; (see the table below)

A – H	I – P	Q – Z
Antigua and Barbuda	Jamaica	St Kitts and Nevis
Australia	New Zealand	St Lucia
The Bahamas		St Vincent and the Grenadines
Barbados		Trinidad and Tobago
Belize		United States of America
Canada		
Dominica		
Grenada		
Guyana		

Source: UKVI Website, p.36 of
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/303848/T4_Guidance_06-04-14.pdf
[Accessed 10/06/2016]

Appendix 3: Examples of 'Country Specific English Language Tests' (CSELTS)

Adapted from the website of Swansea University, this is an example of the alternative qualifications from outer circle countries that universities accept as proof of English (other countries listed on the website are Norway, Sweden, Netherlands, Belgium, France, Denmark, Austria, Switzerland and Finland, but not listed below):

"Taught via the medium of English"

"English Language information: You will not be required to take a secure English language test if you have been taught in English at secondary level or above in one of the following countries and have obtained the qualification and minimum grade listed."

Country	Qualification and Grade
Bermuda	Secondary School Certificate - Grade C or above in English
Botswana	General Certificate of Secondary Education - Grade C or above in English
Brunei	Cambridge Ordinary Level - Grade C or above in English
Cameroon	Ordinary Level - Grade C or above in English
Gambia	WAEC - Grade C6 or above in English
Ghana	WAEC - C6 or above in English
Hong Kong	HKDSE - Grade 3 or above in English
India	Standard XII in English Undergraduate 65%+ Postgraduate 70%+
Kenya	KCSE - Grade D or above in English
Malawi	MCSE – Grade C6 or above in English
Malaysia	SPM – Grade 6 or above in English
Malta	Second Education Certificate - Grade C or above in English
Mauritius	Ordinary Level – Grade C or above in English
Nigeria	WAEC – Grade 6 or above in English
Sierra Leone	WAEC – Grade C6 or above in English
Singapore	Ordinary Level – Grade C or above in English
South Africa	National Senior Certificate – Grade C/5 or above in English
Tanzania	Certificate of Secondary Education - Grade C or above in English
Uganda	UCE – Grade C or above in English
Zambia	Zambia School Certificate – Grade 6 or above in English
Zimbabwe	CSE – Grade 6 or above in English

Source: http://www.swansea.ac.uk/postgraduate/apply/taughtviathemediumofenglish/ [Accessed 10/06/2016]

Appendix 4: The Alignment of the PTE Academic Score Ranges with the CEFR

Use the table below to find out what PTE Academic scores in the range from 10 to 84 (A1 to C1) mean. The table shows the score ranges that predict some degree of performance at the next level, and it explains what you are likely to be able to do within those score ranges.

PTE Academic Score	Common European Framework Level	Level Descriptor © Council of Europe	What does this mean for a score user?
76 - 84	C1	Can understand a wide range of demanding, longer texts and recognise implicit meaning. Can express him/herself fluently and spontaneously without much obvious searching for expressions. Can use language flexibly and effectively for social, academic and professional purposes. Can produce clear, well-structured, detailed text on complex subjects, showing controlled use of organisational patterns, connectors and cohesive devices.	C1 is a level at which a student can comfortably participate in all post-graduate activities including teaching. It is not required for students entering university at undergraduate level. Most international students who enter university at a B2 level would acquire a level close to or at C1 after living in the country for several years, and actively participating in all language activities encountered at university.
59 - 75	B2	Can understand the main ideas of complex text on both concrete and abstract topics, including technical discussions in his/her field of specialisation. Can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without strain for either party. Can produce clear, detailed text on a wide range of subjects and explain a viewpoint on a topical issue giving the advantages and disadvantages of various options.	B2 was designed as the level required to participate independently in higher level language interaction. It is typically the level required to be able to follow academic level instruction and to participate in academic education, including both coursework and student life.
51 – 58	Predicts success on easiest tasks at B2	Has sufficient command of the language to deal with most familiar situations, but will often require repetition and make many mistakes. Can deal with standard spoken language, but will have problems in noisy circumstances. Can exchange factual information on familiar routine and non-routine matters within his/her field with some confidence. Can pass on a detailed piece of information reliably. Can understand the information content of the majority of recorded or broadcast material on topics of personal interest delivered in clear standard speech.	
43 - 58	B1	Can understand the main points of clear standard input on familiar matters regularly encountered in work, school, leisure, etc. Can deal with most situations likely to arise whilst in an area where the language is spoken. Can produce simple connected text on topics, which are familiar or of personal interest. Can describe experiences and events, dreams, hopes and ambitions and briefly give reasons and explanations for opinions and plans.	B1 is insufficient for full academic level participation in language activities. A student at this level could 'get by' in everyday situations independently. To be successful in communication in university settings, additional English language courses are required.
36 – 42	Predicts success on easiest tasks at B1	Has limited command of language, but it is sufficient in most familiar situations provided language is simple and clear. May be able to deal with less routine situations on public transport e.g., asking another passenger where to get off for an unfamiliar destination. Can re-tell short written passages in a simple fashion using the wording and ordering of the original text. Can use simple techniques to start, maintain or end a short conversation. Can tell a story or describe something in a simple list of points.	
30 - 42	A2	Can understand sentences and frequently used expressions related to areas of most immediate relevance (e.g., very basic personal and family information, shopping, local geography, employment). Can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters. Can describe in simple terms aspects of his/her background, immediate environment and matters in areas of immediate need.	A2 is an insufficient level for academic level participation.
10 - 29	A1 or below	Can understand and use familiar everyday expressions and very basic phrases aimed at the satisfaction of needs of a concrete type. Can introduce him/herself and others and can ask and answer questions about personal details such as where he/she lives, people he/she knows and things he/she has. Can interact in a simple way provided the other person talks slowly and clearly and is prepared to help.	A1 is an insufficient level for academic level participation.

Source: Pearson (2012a, p.4), Interpreting the Academic Score Report, (available from the PTE Academic website as a PDF file)

Appendix 5: Test Specification for PTE Academic

Purpose of the instrument

The test will measure English language proficiency for communication in English-medium tertiary setting.

Construct or domain that will be measured

Communicative language skills will be assessed for reception, production and interaction in the oral and written modes as these skills are needed to successfully follow courses and actively participate in education and training where English is the language of instruction.

Framework of the instrument

- Presentation mode: oral-audio or video, graphic, or any combination thereof;
- Content of stimulus: such as instruction, definition, explanation, description, argument
- Task type: such as retrieving information, interpreting, deducting, combining information from different sources, evaluating
- Required competences: such as grammatical, lexical, phonological, pragmatic, strategic
- Response mode: oral or written
- Response format: such as multiple choice (single or multiple answers), clicking hotspots in texts, selecting from drop down lists, drag and drop, highlight, short answers, extended response

Text length

The test will take up to 3 hours.

Context in which the instrument is to be used

English-medium education and training and related professional fields

Characteristics of intended participants

Learners of English as a second or other language who are applying for admission to courses where English is the language of instruction or admission to professional bodies or skilled professions

Psychometric properties

The items will all have a demonstrable and robust relationship with the reporting scales towards which they purport to contribute. Scores on reporting scales will have a standard error of measurement no greater than 15 percent of the score range covering 68% of the intended target population

Conditions and procedure of administering the instrument

Tests will be administered on computer in dedicated Pearson test centers

Procedures of scoring

Depending on the item format, some items will be scored dichotomously, but the majority of items will be scored polytomously. Item scores for closed response will be generated by machine, whereas item scores for constructed, open responses will be generated by automatic scoring system trained on initial human ratings.

Reporting of the results

Overall, communicative skills and enabling skills scores will be reported as a profile of the candidate's level of ability. In addition scores will be provided in numerical form and in relation to the CEF.

Source: Zheng, Y. & De Jong, J. (2011, p.4) Research Note: Establishing Construct and Concurrent Validity of Pearson Test of English Academic

Appendix 6. A Description of the PTE Academic "Enabling Skills"

Grammar	Correct use of language with respect to word form and word order at the sentence level
Oral fluency	Smooth, effortless and natural-paced delivery of speech
Pronunciation	Production of speech sounds in a way that is easily understandable to most regular speakers of the language. Regional or national varieties of English pronunciation are considered correct to the degree that they are easily understandable to most regular speakers of the language
Spelling	Writing of words according to the spelling rules of the language. All national variations are considered correct, but one spelling convention should be used consistently in a given response
Vocabulary	Appropriate choice of words used to express meaning, as well as lexical range
Written discourse	Correct and communicatively efficient production of written language at the textual level. Written discourse skills are represented in the structure of a written text, its internal coherence, logical development and the range of linguistic resources used to express meaning precisely

Source: Pearson (2012a, p.3) The PTE Academic Score Guide, Pearson Education Ltd

Available from:

http://www.pearsonlongman.com/examsplace/PTE_Academic/index.html [Accessed 10/06/2016]

Appendix 7: Extract from Self-Rating Tool & Recruitment Survey

Below is an extract from the online survey with the full version accessible via the UEL hyperlink below:

URL Link:
http://www2.warwick.ac.uk/study/csde/gsp/eportfolio/directory/pg/live/elrmag/studentquestionnairemainstudy/ [Accessed 10/06/2016]

Example of the self-rating questionnaire for Listening Skills:

20) ENGLISH LISTENING SKILLS: Please Tick ONE sentence that best describes your current reading ability *

○ 1) I can understand enough to be able to meet needs of a concrete type provided speech is clearly and slowly articulated.

○ 2) I can understand the main points of clear standard input on familiar matters regularly encountered in work, school, leisure, etc.

○ 3) I can understand the information content of the majority of recorded or broadcast material on topics of personal interest delivered in clear standard speech. I can deal with standard spoken language, but will have problems in noisy circumstances

◉ 4) I can understand extended speech and lectures and follow even complex lines of argument provided the topic is reasonably familiar.

○ 5) I can understand extended speech even when it is not clearly structured and when relationships are only implied and not signalled explicitly. I can understand television programmes and films without too much effort.

Appendix 8: 'Experiencing Academia' as an interview tool

'Experiencing Academia'

'Experiencing Academia' is my concept derived from the academic situations, problems and tasks in previous studies. I used these in my questionnaire and interview tools. Here, for convenience, I have named these three terms as 'aspects' (of performance). Four studies were used as sources for these and the aspects of performance taken from each study are displayed below:

1. Byrne & Flood (2005)

All nine aspects below were taken from Table 3 (2005:117) - "Preparedness for Higher Education":

- Work independently
- Able to initiate your own study activities
- Planning your study time
- Take responsibility for your own learning
- Willing to ask for help from your lecturers/tutors
- Comfortable working in groups
- Ability to complete written assignments
- Willing to participate in class
- Able to evaluate your own progress

2. Woodrow (2006)

All 11 aspects below were taken from Appendix B: Student Questionnaire Items (2006:69-70, Question 4 – in one table):

- Finding resources
- Understanding language in textbooks and readings
- Understanding information in textbooks and readings
- Understanding language of instructors
- Understanding content of lectures or tutorials
- Understanding language of other students
- Taking notes in lectures
- Taking part in class discussion
- Asking questions
- Preparing assignments – content
- Preparing assignments – language

3. Ferguson & White (1994) 'List of Appendices'

All ten aspects below were taken from pages 14-15 of the appendices to their study (1994). These were sent to the researcher from personal correspondence with Ferguson. They are from Questionnaire 4. Section 2 'Problems' and page 17 of the same appendices, Questionnaire 5 – Table 1 'Factors Affecting Success':

- Overall difficulty in understanding the subject matter
- Difficulty in completing the set work
- Too great a quantity of materials to be studied
- Inefficient study methods
- Unsatisfactory teaching methods
- Difficulty in obtaining books
- Irrelevance of the subject matter
- Background knowledge/previous experience of the subject
- Self-confidence/belief in your own ability
- Well-taught/well-organised course

4. Banerjee (2003:418)

All ten aspects were taken from Appendix 3A - Student Self-Assessment Questionnaire (1), PART ONE Table on p.418:

- Listening to and understanding lectures
- Listening to and understanding the media (TV, radio)
- Understanding people in shops/on the street/in public places
- Participating in seminars and tutorials
- Participating in conversations on non-academic topics in non-academic settings (cafes, college canteens, etc)
- Reading academic books, manuals, journals
- Reading for pleasure and relaxation (novels, short stories, poetry)
- Reading for general interest and information (newspapers, magazines)
- Taking notes from textbooks
- Writing for non-academic purposes (letters, reports)

5. Added Aspects

These aspects were added as it was felt they were missing in the tools of the above studies:

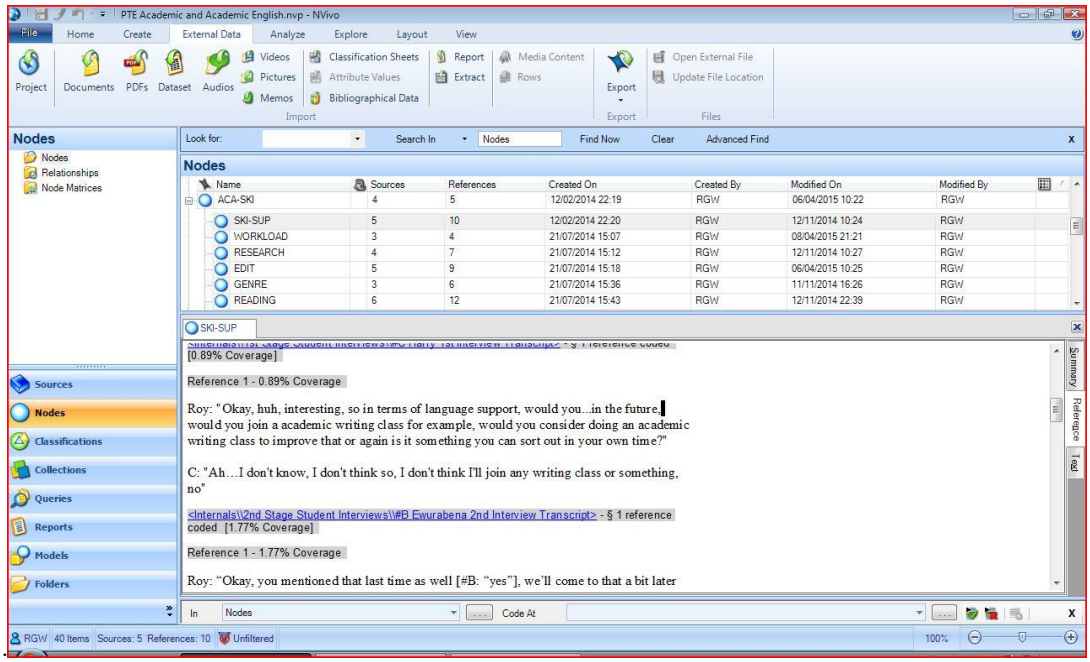
- Using I.T.
- UK Academic Culture
- Doing presentations
- Answering questions from audience after presentations

Appendix 9: Interview Transcription Conventions

A, B, etc…	- denotes the participant speaking
Interviewer:	- denotes the researcher speaking
… (Ellipsis)	- denotes a pause or hesitation in speech
Erm	- denotes a natural hesitation or pause in searching for words
Uhuh	- indication of understanding/sign to continue/ "I'm listening" etc.
(Laughs)	- denotes non-verbal behaviour (laughter for example)
[Inaudible]	- denotes an utterance than cannot be heard clearly on the digital recording/cannot be recalled.

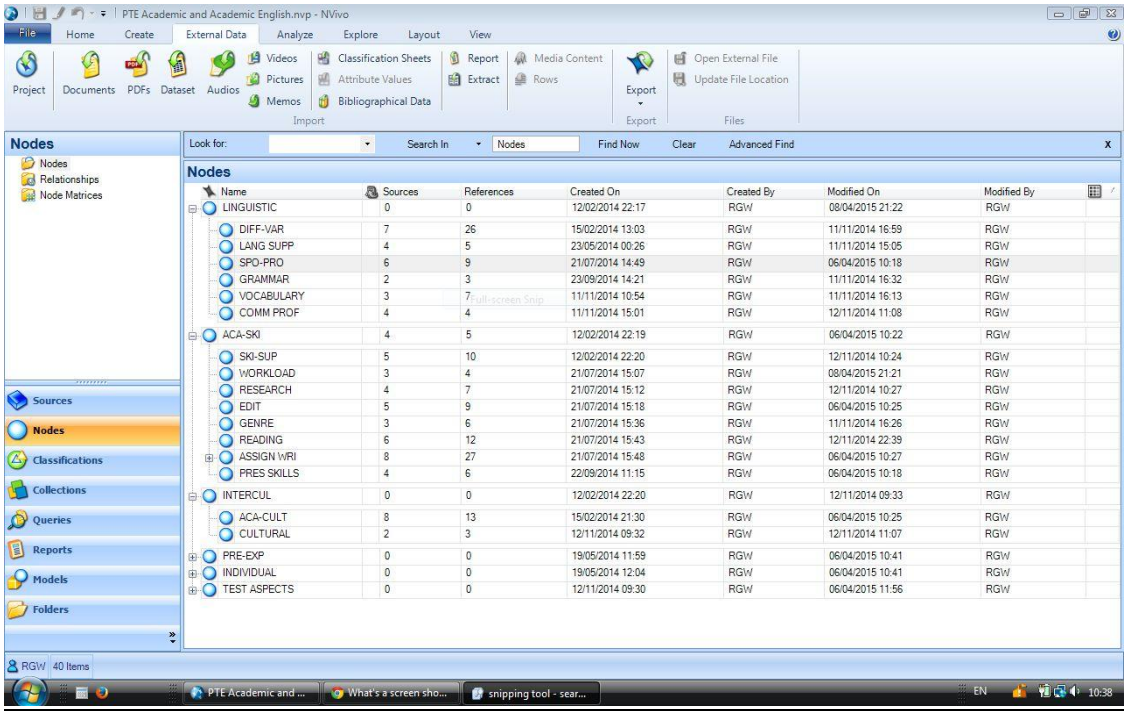
Appendix 10: Nvivo Software Screen Shot – Nodes and Data

A screen shot illustrating the empirical data in my coding procedure. The large window at the bottom shows the actual chunk of data (dialogue or utterance) that is assigned to the child node. In this case the utterance is about "Skills Support" which itself is linked to the wider concept of "Academic skills" (parent node).



Appendix 11: Example of Multiple Parent and Child Nodes – Nvivo Screen Shot

A screen shot from Nvivo showing the 'tree of nodes' e.g. in the first column is the name of the parent node e.g.' linguistic', then sub-headings representing the child nodes under that node: Examples under 'linguistic' include: "Different Varieties", "Language Support", "Spoken Production", "Grammar", "Vocabulary" and "Communicative Proficiency". In the second column is an indication of how many sources contribute to the code and then in the third column, the number of actual data chunks from the tutorials, interviews or surveys that constitute the data.



Appendix 12: New Parent Nodes and Definitions

1. LINGUISTIC
"Any mention of a particular problem related to language proficiency - e.g. grammar, pronunciation, vocabulary preventing progress on the course. The student may mention a deficiency or shortfall in their language proficiency which causes problems on the course"

2. ACA-SKI (Academic Skills)
"The student mentions a problem with a particular skill needed in academia for example awareness of the format of essay writing, referencing and other conventions or any other academic skills, for example the ability to read widely and quickly or preparing and conducting presentations"

3. INTERCULT (Intercultural)
"A mention of cultural differences impeding their performance - for example, informality in dealing with tutors, (power distance differences between cultures) feelings of not being able to question tutors or even ask for help, difficulties in working with other students etc…"

4. PRE-EXP (Previous experience)
"A factor influencing academic performance: Previous experience including academic (educational background, medium of instruction), and professional experience. Ownership of the language fits here in relation to the experiences of the test takers from their background in the outer circle"

5. INDIVIDUAL
"A variable that expresses the individual character or personality of the student as affecting performance such as personality, ability to be pro-active, confidence or lack of confidence and dissatisfaction with an aspect of their studies"

6. TEST ASPECTS
"Comments on aspects of the test itself, comments on preparation, performance, use of the test scores, the politics of the test, views on the format of the test, and difficulties or dissatisfaction that test takers had with the test itself"

Appendix 13: Example of Consent Form - Case Study Participants

Case Study - CONSENT FORM

Project Title: The PTE Academic Score Report and Student Performance at University

This ESRC-Funded Collaborative PhD (see information sheet) will examine the relationship between your language skills profiles from your Pearson Test of English Academic and your performance at university as perceived by yourself, and with reference to your EAP tutors and academic lecturers. My study should help you to reflect on your test experience and academic life at university and allow you to assess how you could progress in the future. In particular, I will examine how well you feel you were prepared for academic study.

Name of Researcher: Roy Wilson, PhD Student of The University of Warwick

I confirm that I have read and understood the information sheet dated …….. for the above project which I may keep for my records and have had the opportunity to ask any questions I may have. I agree to take part in the above study and am willing to…(Please tick if you consent):

☐ …complete an online questionnaire.
☐ …be interviewed for up to an hour.
☐ …allow the interview to be recorded and transcribed.
☐ …give permission to the researcher to look at my PTE Ac. test scores, university assignment marks and written academic work including comments from tutors.

Note: All identities will remain anonymous; your name and university affiliation will not be released publicly.

I understand that my information will be held and processed for the following purposes in which my identity will remain anonymous at all times:

* Inclusion in the researcher's PhD study at The University of Warwick.
* Possible publication in an academic journal.
* Shared with the organisation, Pearson Education (who have signed an ethics agreement)
* Included in presentations and talks that the researcher may give.

I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason, without being penalised or disadvantaged in any way. My identity and that of my institution will remain anonymous at all times.

A) Name of Participant

Date Signature

B) Name of person taking consent if different from researcher

Date Signature

C) Researcher

Date Signature

Appendix 14: Survey of UK University Admissions

Grouping of Universities According to Admissions Policies on CSELTs and SELTs			
Accepting and displaying a list of OC alternatives (n=23)	Accepting of a limited number of OC nations or with conditions/caveats (n=33)	Not accepting any alternatives to SELTs (n=45)	Those stating alternatives may be acceptable on email enquiry (n=31)
Bath Spa	Abertay	Aberdeen	No Response (2)
Bournemouth	Aberwystwyth	Anglia Ruskin	Queen Margaret (Edin)
Brighton	Bath	Birkbeck	West London
Bristol	Bedfordshire	Brunel	
Cardiff	Birmingham	Buckinghamshire New	Yes – Positive Response (20)
Central Lancashire	Birmingham City	Canterbury-Christchurch	Aston
Edinburgh Napier	Buckingham	Cardiff Metropolitan	Bangor
Exeter	Cambridge	Chichester	Bolton
Glasgow Caledonian	Chester	City	Bradford
Gloucestershire	Coventry	Cranfield	De Montfort
Lancaster	Cumbria	Dundee	Derby
Leicester	Durham	Edge Hill	Glasgow
Lincoln	Essex	Edinburgh	Greenwich
Loughborough	Hertfordshire	Glyndwr	Keele
Manchester-Metropolitan	Hull	Goldsmiths	Leeds
Middlesex	Imperial College	Guildhall School of-Music and Drama	Leeds Beckett
Nottingham	Liverpool	Harper Adams	Liverpool John Moores
Nottingham Trent	Liverpool Hope	Herriott-Watt	Manchester
Oxford	Newcastle	Heythrop College	Newman
Sheffield	Oxford Brookes	Huddersfield	Northumbria
Swansea	Portsmouth	Kings College London	Reading
Uni of East Anglia	Plymouth	Kingston	St Mark and St John
Worcester	Queens Belfast	Leeds Trinity	Strathclyde
	Robert Gordon	London Metropolitan	Sunderland
	Roehampton	LSE	UEL
	St Andrews	London School of -Hygene & Tropical-Medicine	CSELT not accepted SELT only (9)
	Salford	Northampton	Arts Uni Bournemouth
	Southampton	Regent's University	Bishop Grosseteste
	Staffordshire	Royal Holloway	Falmouth
	Sussex	Sheffield Hallam	Glasgow School of Art
	UWE	SOAS	Kent
	Warwick	South Wales	Queen Mary
	Wolverhampton	Southampton Solent	St Georges
		Stirling	St Mary's Twickenham
		Surrey	Southbank
		Teeside	
		Trinity St David	
		UAL	
		UCL	
		Ulster	
		West of Scotland	
		Westminster	
		Winchester	
		York	
		York St John	

Appendix 15: Outer Circle Survey Tool

Selected examples of questions from the online survey can be seen below. The full survey can be viewed online at the following link:

http://www2.warwick.ac.uk/study/csde/gsp/eportfolio/directory/pg/live/elrmag/outercirclesurvey/ [Accessed 28/07/2016]

14) Which of these methods were used as a way of proving your English language proficiency when you applied to your chosen university? (tick all that apply) *

- A high school English certificate from your country
- An international Academic English test e.g. IELTS, PTE, TOEFL etc…
- Undergraduate degree from your country in English
- A face-to-face interview
- A Skype interview
- A written assignment
- a language test offered by the university you applied to
- Other (please describe below)

Other Example Questions

19) Before coming to the UK, to what extent did you have a level of English that was appropriate for your UK university course?

20) To what extent have you encountered any difficulties to do with English language proficiency on your UK university course?

21) To what extent have you encountered any difficulties to do with academic skills on your UK university course?

Appendix 16: Interview Schedule for Pearson Representatives

(Country X = Nigeria, India & Malaysia)

Part 1. The Students (PTE Ac. Test takers in country X)

1.1 What would be the profile of a typical test-taker from your country in terms of:

- Which countries they apply for in terms of university study
- Typical disciplines applied for
- Ages/gender
- Educational background – (from L1 medium schools or English medium)
- What test takers from country X generally score on the PTEA
- Any average overall score records known
- Whether scores are typically high compared to other Commonwealth countries or neighbouring 'expanding circle' countries

1.2 What feedback from test-takers do you get in terms of:

- Students' satisfaction with scores
- Students' attitudes to, and experiences of taking the test

1.2 Is there a difference in PTEA test performance in terms of:

- Those who went to L1-medium schools
- Those who went to English-medium schools

1.4 What is the experience of students from country X in applying for university places abroad in terms of:

- Whether country X High school English exams are accepted
- Whether SELTS are insisted upon

Part 2. Marketing & Policy

2.1 How is the test actually promoted in country X?

- For example, is the test promoted as essential for entry into the UK no matter what the educational background/previous qualifications of the individual?
- Does the UK Embassy/UK Visa and Immigration insist on country X nationals taking a SELT?

2.2 What is the Pearson advice given to students on:

- Tier 4 Visas.
- UK university English language requirements and proof of proficiency
- Use of country X High school English qualifications for university admissions purposes

2.3 University Policies

- What is the interviewee's understanding of a university's English language requirements policy at admissions?
- Do the universities have to say 'do a SELT'?
- Are they free to admit students with any qualification deemed suitable/matching B2 by the university?

Part 3. Other

3.1 If a country X student comes to you and says "I have an A-C grade in the domestic exams in English, why should I do a PTE Academic" what would you say to them?

3.2 Some country X students may consider themselves native speakers of English or very close to that, so in that case:

- How would they demonstrate their proficiency without having to take a SELT?
- Do these individuals still have to take either IELTS or PTEA in order to get student visas (as nationals of country X)?
- If so, what are their attitudes to having to take the test?

3.3 In your opinion, what does the PTEA provide or 'say' about the language proficiency of a student that their high school qualification in English does not?

3.4 Is PTEA a "better" qualification for academic English than students' country X high school qualification in terms of:

- More internationally recognised? (RECOGNITION)
- More comprehensive measure of their English? (FOUR SKILLS)
- Considered a 'better' or 'more accurate' indicator of their English level because it is based on a corpus (PICAE)

3.5 Any other questions as appropriate in the interview

Appendix 17: Self-Rating of Proficiency & CEFR Levels in Recruitment Survey

Note: The 'actual PTEA scores' are all self-reported PTEA scores.

Below are all the tables illustrating individual test taker information for self-rating of proficiency for the overall score and four sub-scores (using the PTEA alignment descriptors with the CEFR) compared against the CEFR level according to PTEA alignment with the CEFR (Pearson 2012a, p.40-41).

Nationality Key: AUS = Australia; BRA = Brazil; CAM = Cameroon; CAN = Canada; COL = Colombia; DEN = Denmark; GHA = Ghana; IND = India; IRQ = Iraq; ISR = Israel; JPN = Japan; KEN = Kenya; KUW = Kuwait; LIB = Libya; NIG = Nigeria; PAK = Pakistan; ROM = Romania; S.A. = South Africa; S.SUD =South Sudan; TUR = Turkey;

Colour Coding Key:
Yellow highlight = the four cases who took part in the case study tutorials
Black = individuals who under-rated themselves compared to their PTEA score
Green = individuals whose self-rating matched with their PTEA proficiency level
Red = individuals who over-rated themselves compared to their PTEA score

Nb: Student #23 from India did not report any PTEA scores so was omitted from the tables altogether. Student #8 from Russia reported an overall PTEA score but did not report any communicative skills scores so is omitted from table 3 onwards.

Table 77 Actual Overall PTEA & CEFR Level vs. Students' Self-Rating of CEFR Level

Overall		Actual Overall PTEA Scores (equivalent CEFR Band)					
		A1 PTE 10-29	A2 PTE 30-42	B1 PTE 43-50	B1+ PTE 51-58	B2 PTE 59-75	C1-C2 PTE 76+
Students' Self Rating of CEFR Level	A1						
	A2			(#22 IND)			
	B1		(#29 TUR)				
	B1+		(#43 S.SUD)		(#3 JPN)	(#45 NIG)	
	B2		(#6 IND)	(#1 NIG), (#21 IND) (#33 PAK) (#47 NIG)	(#18 COL) (#20 IND) (#42 NIG)	(#8 RUS) (#13 LIB) (#24 NIG) (#26 PAK) (#31 ISR) (#40 IND)	(#2 ROM)
	C1-C2		(#10 NIG)	(#32 IRQ), (#34 KUW)	(#7 NIG) (#9 NIG) (#19 TUR) (#25 CAM) (#39 BRA) (#41 NIG)	(#5 NIG) (#11 IND) (#15 AUS) (#30 IND) (#35 PAK) (#36 NIG) (#44 GHA) (#46 NIG)	(#4 IND) (#12 KEN) (#14 IND) (#16 IND) (#17 S.A.) (#27 IND) (#28 KEN) (#37 CAN) (#38 DEN)

Table 78 Actual Listening PTEA & CEFR Level vs. Students' Self-Rating of CEFR Level

Listening		Actual Listening PTEA Scores (equivalent CEFR Band)					
		A1 PTE 10-29	A2 PTE 30-42	B1 PTE 43-50	B1+ PTE 51-58	B2 PTE 59-75	C1-C2 PTE 76+
Students' Self Rating of CEFR Level	A1						
	A2			(#43 S.SUD)		(#25 CAM)	
	B1			(#9 NIG)	(#3 JPN)	(#24 NIG)	
	B1+		(#29 TUR)	(#22 IND)	(#47 NIG)	(#40 IND)	
	B2	(#2 ROM)	(#6 IND)	(#1 NIG) (#18 COL) (#21 IND)	(#20 IND)	(#5 NIG) (#11 IND) (#19 TUR) (#42 NIG) (#31 ISR)	(#14 IND)
	C1-C2			(#10 NIG) (#32 IRQ) (#34 KUW)	(#33 PAK) (#39 BRA) (#41 NIG)	(#7 NIG) (#13 LIB) (#15 AUS) (#26 PAK) (#44 GHA) (#45 NIG) (#46 NIG)	(#4 IND) (#12 KEN) (#16 IND) (#17 S.A.) (#27 IND) (#28 KEN) (#30 IND) (#35 PAK) (#36 NIG) (#37 CAN) (#38 DEN)

Table 4 Actual Reading PTEA & CEFR Level vs. Students' Self-Rating of CEFR Level

Reading		Actual Reading PTEA Scores (equivalent CEFR Band)					
		A1 PTE 10-29	A2 PTE 30-42	B1 PTE 43-50	B1+ PTE 51-58	B2 PTE 59-75	C1-C2 PTE 76+
Students' Self Rating of CEFR Level	A1						
	A2		(#22 IND) (#43 S.SUD)				
	B1	(#6 IND)	(#21 IND)				
	B1+	(#2 ROM) (#10 NIG) (#47 NIG)	(#1 NIG) (#25 CAM) (#29 TUR) (#32 IRQ) (#34 KUW) (#33 PAK) (#42 NIG)	(#20 IND) (#26 PAK) (#40 IND)	(#3 JAP) (#45 NIG)	(#4 IND) (#5 NIG) (#9 NIG) (#24 NIG)	
	B2			(#19 TUR)	(#11 IND)	(#31 ISR) (#35 PAK) (#46 NIG)	(#16 IND) (#27 IND)
	C1-C2			(#41 NIG)	(#18 COL) (#39 BRA)	(#7 NIG) (#12 KEN) (#13 LIB) (#15 AUS) (#17 S.A.) (#30 IND) (#44 GHA) (#36 NIG)	(#14 IND) (#28 KEN) (#37 CAN) (#38 DAN)

Table 5 Actual Speaking PTEA & CEFR Level vs. Students' Self-Rating of CEFR Level

Speaking		Actual Speaking PTEA Scores (equivalent CEFR Band)					
		A1 PTE 10-29	A2 PTE 30-42	B1 PTE 43-50	B1+ PTE 51-58	B2 PTE 59-75	C1-C2 PTE 76+
Students' Self Rating of CEFR Level	A1						
	A2			(#6 IND) (#22 IND) (#43 S.SUD)			
	B1	(#29 TUR)	(#34 KUW)	(#1 NIG) (#3 JPN)		(#21 IND) (#45 NIG)	
	B1+	(#2 ROM)		(#20 IND)	(#25 CAM) (#42 NIG)	(#9 NIG) (#33 PAK) (#40 IND)	(#11 IND)
	B2				(#10 NIG) (#18 COL) (#32 IRQ)	(#5 NIG) (#7 NIG) (#13 LIB) (#15 AUS) (#24 NIG) (#26 PAK) (#31 ISR) (#44 GHA)	(#14 IND) (#16 IND)
	C1-C2		(#39 BRA)	(#41 NIG) (#47 NIG)		(#19 TUR)	(#4 IND) (#12 KEN) (#17 S.A.) (#27 IND) (#28 KEN) (#30 IND) (#35 PAK) (#36 NIG) (#37 CAN) (#38 DEN) (#46 NIG)

Table 6 Actual Writing PTEA & CEFR Level vs. Students' Self Rating of CEFR Level

Writing		Actual Writing PTEA Scores (equivalent CEFR Band)					
		A1 PTE 10-29	A2 PTE 30-42	B1 PTE 43-50	B1+ PTE 51-58	B2 PTE 59-75	C1-C2 PTE 76+
Students' Self Rating of CEFR Level	A1						
	A2	(#10 NIG)	(#43 S.SUD)	(#9 NIG) (#22 IND)			(#37 CAN)
	B1	(#6 IND)			(#11 IND) (#45 NIG)	(#24 NIG) (#40 IND)	
	B1+	(#2 ROM)	(#21 IND)	(#19 TUR) (#29 TUR) (#32 IRQ)	(#3 JAP)	(#5 NIG) (#13 LIB) (#20 IND)	
	B2			(#1 NIG) (#33 PAK) (#47 NIG)	(#15 AUS) (#18 COL)	(#25 CAM) (#26 PAK) (#31 ISR) (#34 KUW) (#35 PAK) (#39 BRA) (#42 NIG)	(#4 IND) (#16 IND) (#27 IND) (#36 NIG)
	C1-C2					(#7 NIG) (#30 IND) (#41 NIG) (#44 GHA) (#46 NIG)	(#12 KEN) (#14 IND) (#17 S.A.) (#28 KEN) (#38 DAN)

Appendix 18: First Student Tutorial Protocol

FIRST SECTION

Introductory comments, offer thanks, explain purposes of tutorials, sign consent form, inform of length of interview, explain presence of recording tool (electronic dictaphone), offer book voucher as a thank you, and check original Pearson Score Profile (see original scores if student has them and check all ten sub-scores).

Their PTEA Score Profile:
Follow-up from student's responses on the questionnaire (student specific).
Ask about self-rating of proficiency choices.

- How satisfied did/do you feel with your score profile?

- Did your scores give you expectations about what you might find difficult or easy about university life? For example, does a high speaking score give you confidence in doing presentations in English, speaking to tutors and academic staff, asking questions in class etc?

- Did all your scores give you confidence in being able to cope with all aspects of university life?

- Were there any scores that did not give you confidence?

- What meaning/use does the score profile have for you now that you are at university?

SECOND SECTION

Previous Experience & Training - Questions from:
"Background knowledge/previous experience of the subject"
(Ferguson & White, 1994b:17) (Woodrow (2006:56)

- What previous education do you have in English?
- Do you have professional experience in the subject?
- Was this experience using English?
- Do you think this experience has been useful for your current studies? – Why?/Why not?

THIRD SECTION – Actual experience

Language Support Needed & Offered
For specific question-wording See: O'Loughlin, (2008:51)

- Was support required before and since starting the course? (Instigated by university)
- Did the student feel that they needed support?
- Did the student seek support?
- Are they accessing support now?

Students Evaluating their 'Actual Experience'

Distribute "actual experience cards" and explain the exercise rationale (student reads the card, decides to what extent this is an issue for them and discuss (don't spend too long on each card) WERE THESE ISSUES DUE TO LANGUAGE PROBLEMS?

FORTH SECTION

Assessment & Assignments

- Have they done any assignments/assessments?
- What was the feedback?
- What will they be expected to do on their course?
- Ask them to bring along some written work and feedback in the 2nd term

Access to written work and tutors

- Will I be able to access their work/comments from tutors?
- Will I be able to access their tutors?
- For Tutorials 2 & 3, in the 2nd and 3rd terms, students fill in the self-assessment form (CEFR) before the interview to discuss in the Tutorial.
- Ask about further participation in interviews in terms two and three.

Appendix 19: Second Student Tutorial Pre-interview Online Questionnaire

These questions were adopted from Banerjee (2003, p.427). They are part of the full survey which can be viewed at the link below:

URL Link:
http://www2.warwick.ac.uk/study/csde/gsp/eportfolio/directory/pg/live/elrmag/2ndself-assessment/ [Accessed 10/06/2016]

7) How much difficulty have you had with your written coursework and exams?	very often	often	sometimes	rarely	never
Understanding the purpose of the assignment	○	○	○	○	○
Writing a relevant answer	○	○	○	○	○
Organising my ideas in an appropriate structure	○	○	○	○	○
Expressing my ideas clearly	○	○	○	○	○
Writing in an acceptable style	○	○	○	○	○
Meeting the word limit	○	○	○	○	○
Finding enough to write (content)	○	○	○	○	○
Vocabulary problems	○	○	○	○	○
Grammar Problems	○	○	○	○	○
Spelling Problems	○	○	○	○	○
Punctuation problems	○	○	○	○	○
Using what I have read in my written work	○	○	○	○	○
Other (Please describe in box below)	○	○	○	○	○

Appendix 20: Example of University Marking Criteria (Case C)

Appendix 1 –University Generic Marking Criteria

These generic marking criteria will help you to distinguish between the different levels of performance that are required for the different grades applied to assessments. You should use them to guide your work when you are not provided with more specific criteria. In particular, you should notice that the higher grades place more emphasis on skills such as independent reading and the correct use of these sources, critical analysis, concise writing and well presented arguments. On the other hand, the lower grades are obtained when the work is incoherent and descriptive.

Band	**FHEQ level 7 (Level M) Criteria**
90-100%	An excellent critical and complete demonstration of understanding in all key areas of knowledge relevant to the work and demonstrating an innovative and creative approach. Evidence throughout the work of a sustained ability to synthesise and interpret complex concepts, to make inferences and to provide an original and/or compelling argument and discussion. Excellent structure and immaculate presentation, with cogent use of academic language and grounded in a pertinent and substantial selection of source materials. Excellent use of appropriate analytical and research methods and addresses ethical considerations in an informed and perceptive manner. Exceptional ability to link and critically analyse theory and practice where appropriate.
80-89%	An excellent, critical and systematic demonstration of understanding in all key areas of knowledge relevant to the work. Evidence throughout of the ability to synthesise and interpret complex concepts to provide a compelling argument and discussion. Very good structure and presentation, with confident use of academic language and grounded in a relevant and extensive selection of source materials. Excellent use of appropriate analytical and research methods and fully addresses ethical considerations. Excellent ability to link and critically analyse theory and practice where appropriate.
70-79%	An excellent, critical and organised demonstration of understanding in all key areas of knowledge relevant to the work. Evidence throughout of the ability to synthesise and interpret diverse concepts to provide a sound argument and discussion. Good structure and presentation, with fluent use of academic language and grounded in an appropriate and comprehensive selection of source materials. Very effective use of appropriate analytical and research methods and consideration of ethical implications. Very good ability to link and critically analyse theory and practice where appropriate.
60-69%	A proficient, clearly stated and analytical demonstration of understanding in all key areas of knowledge relevant to the work. Evidence of the ability to integrate and analyse diverse concepts in a rational and logical argument and discussion. Well structured and clearly presented work, with fluent use of academic language and utilising a relevant and extensive range of source materials. Effective use of appropriate analytical and research methods and consideration of ethical issues. Good ability to link and critically analyse theory and practice where appropriate.
50-59%	An acceptable and substantiated demonstration of understanding in all key areas of knowledge relevant to the work. Evidence of the ability to integrate and analyse diverse concepts in a reasoned and valid argument and discussion. Adequately structured and presented work, with clear use of academic language and

	reference to a sufficient range of relevant source materials. Adequate use of appropriate analytical and research methods and does address ethical considerations. Effective linking of theory and practice where appropriate.
40-49%	A limited, insufficient and/or inaccurate understanding in key areas of knowledge relevant to the work. Insufficient evidence of ability to integrate and analyse concepts to provide a valid discussion. Unacceptably structured and presented work, with insufficient use of academic language and conventions. A limited range of source materials is used. Limited or ineffective use of analytical and research methods and limited coverage of ethical considerations. Inadequate linking of theory and practice where applicable.
30-39%	A descriptive and/or narrative account, with little critical and/or flawed understanding of key areas of knowledge relevant to the work. Insufficient evidence of ability to discuss fundamental concepts. Unclear and and/or unevidenced argument and discussion. Poorly structured and presented work, with little use of academic language and conventions. A narrow and/or inappropriate range of source materials and analytical and research methods is used. Failure to identify ethical considerations and to link theory and practice where applicable.
0-29%	A weakly descriptive and/or narrative account, with no analytical content and/or significant inaccuracies in understanding of key areas of knowledge relevant to the work. Little or no evidence of research and the ability to discuss fundamental concepts. No awareness of ethical issues. Unclear and unsourced arguments and discussion. Flawed structure and presentation, with negligible attention to academic language or conventions. Some or all source materials are unreferenced and/or irrelevant. Failure to link theory and practice where applicable.

Source: remains anonymous so as not to identify the university